



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Durational correlates of English sublexical constituent structure

**Citation for published version:**

Sugahara, M & Turk, A 2009, 'Durational correlates of English sublexical constituent structure' *Phonology*, vol. 26, no. 03, pp. 477-524. DOI: 10.1017/S0952675709990248

**Digital Object Identifier (DOI):**

[10.1017/S0952675709990248](https://doi.org/10.1017/S0952675709990248)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

*Phonology*

**Publisher Rights Statement:**

With permission. © Cambridge University Press. Sugahara, M., & Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology*, 26(03), 477-524doi: 10.1017/S0952675709990248

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Phonology

<http://journals.cambridge.org/PHO>

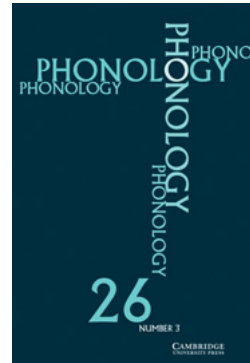
Additional services for **Phonology**:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



---

## Durational correlates of English sublexical constituent structure

Mariko Sugahara and Alice Turk

Phonology / Volume 26 / Issue 03 / December 2009, pp 477 - 524

DOI: 10.1017/S0952675709990248, Published online: 19 November 2009

**Link to this article:** [http://journals.cambridge.org/abstract\\_S0952675709990248](http://journals.cambridge.org/abstract_S0952675709990248)

### How to cite this article:

Mariko Sugahara and Alice Turk (2009). Durational correlates of English sublexical constituent structure. *Phonology*, 26, pp 477-524 doi:10.1017/S0952675709990248

**Request Permissions :** [Click here](#)

# *Durational correlates of English sublexical constituent structure\**

**Mariko Sugahara**

Doshisha University

**Alice Turk**

University of Edinburgh

---

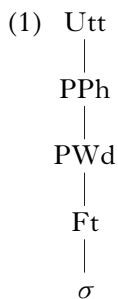
This study investigates whether differences (a) in word-internal morphological structure and (b) in lexical stress patterns are reflected in prosodic constituent structure, by examining duration measurements in Scottish English. In Experiments 1 and 2, at a slow speech rate, stem-final rhymes followed by Level II suffixes were on average 4–6 % longer than corresponding strings in monomorphemic words, and 7–8 % longer than stem-final rhymes followed by Level I suffixes. These results are consistent with the view that stems preceding Level II suffixes are mapped onto prosodic words in the prosodic representation. Experiment 3 obtained no reliable durational differences, even at a slow speech rate, between the initial syllable rhymes of SS words and SW words, which does not provide evidence for the hypothesis that these different stress patterns are represented as differences in foot structure.

---

## 1 Introduction

Current theories of prosodic phonology propose a hierarchy of prosodic constituents which correlate with, but are not necessarily isomorphic to, the morphosyntactic hierarchical structure of stems, words, phrases and sentences (Selkirk 1981, 1986, 1995, Nespor & Vogel 1986, Hayes 1989). The number of levels in the hierarchy, the terms used to refer to them and the way they relate to syntax are matters of debate. Nevertheless, many proposed hierarchies include the constituents shown in (1), i.e. syllable, foot, prosodic word, phonological phrase and utterance.

\* We are grateful to the anonymous reviewers and the associate editor for their helpful comments and suggestions. The experiments reported here were supported by AHRB (Grant No. B-RE-AN9288-APN14615) to the second author, and the pilot study presented in §3.3 was supported by MEXT (Grant No. 18720133) to the first author.



Prosodic constituents serve as the domains of phonological processes (Nespor & Vogel 1986, Selkirk 1986 *et al.*) and also as part of the input to phonetic implementation, where they influence various aspects of the phonetic shape of utterances, including the assignment of segment duration. For example, word-final segments are longer when immediately followed by utterance or PPh boundaries than when utterance-medial or PPh-medial (Cooper & Paccia-Cooper 1980, Beckman & Edwards 1990, Wightman *et al.* 1992, etc.). Segmental duration is also affected by the presence or absence of PWd-level constituent boundaries (Beckman & Edwards 1990, Sproat & Fujimura 1993, Turk & Sawusch 1997, Turk & White 1999, Turk & Shattuck-Hufnagel 2000). For example, Turk & Shattuck-Hufnagel show that /tun/ is longer in *tune acquire*, where it is immediately followed by a PWd boundary, than in *tuna choir*. Sublexical constituent boundaries such as syllable boundaries also show correlations with segmental duration: for example, Redford & Randall's (2005) perception study on English syllabification shows that listeners are more likely to place a syllable boundary at the left edge of consonants when they are longer.

Syllables are not the only sublexical constituents which have been proposed. For example, Aronoff & Sridhar (1983), Szpyra (1989), McCarthy & Prince (1993b) and Cohn & McCarthy (1994) postulate a PWd boundary at a stem-affix boundary within a string of segments corresponding to a single morphological word. However, the existence of such a sublexical PWd boundary has not been tested empirically. Another sublexical prosodic constituent which has been proposed is the word-internal foot. The word-internal foot has been adopted by Selkirk (1980a) and Hayes (1980) to describe stress patterns in English, as well as to explain infixation (McCarthy 1982, Hammond 1999) and children's acquisition of stress patterns (Gerken 1994a, b). However, most of these phenomena can be accounted for without assuming the word-internal foot, and no strong duration-based evidence for the constituent has been obtained so far, as discussed in §3.

The main purpose of this article is to investigate the presence *vs.* absence, and nature, of duration-based evidence for these two sublexical prosodic constituents.

In §2, we present a test of durational evidence for a word-internal PWd boundary, and in §3 a test of durational evidence for word-internal feet. Each of these sections begins with a discussion of theoretical proposals and evidence.

## 2 Level II suffixation and the prosodic word boundary proposal

### 2.1 Level I vs. Level II suffixes

2.1.1 *Two classes of suffixes.* English stems show different phonological behaviour depending on the types of suffix which follow them. Stems followed by Level II suffixes, i.e. inflectional suffixes such as *-ing* (present participle), *-ed*, *-s* and derivational suffixes of Germanic origin such as *-er*, *-ness*, *-less*, *-ing* (gerundive), behave phonologically as if they constitute a word-level constituent on their own: the stems (e.g. *sing* in *singing* and *parent* in *parenting*) and their corresponding independent lexical words have the same surface representations, as shown in (2). In (2a), the stem-final and lexical word-final consonant cluster /ŋg/ is simplified to [ŋ], and in (2b) the stem and the lexical word (henceforth LWd) counterpart share initial stress.

- (2) a. (si/ŋg/ → [ŋ])<sub>Stem</sub> ing      (si/ŋg/ → [ŋ])<sub>LWd</sub>  
       b. ('parent)<sub>Stem</sub> ing              ('parent)<sub>LWd</sub>

In contrast, stems followed by derivational suffixes of Latin origin such as *-al*, i.e. Level I suffixes, show differences in stress placement and segmental behaviour as compared to related independent lexical words.<sup>1</sup> For example, unlike LWd-final consonants, stem-final /ŋg/ consonant sequences do not undergo cluster simplification before Level I suffixes, as shown in (3).

- (3) (diph'tho[ŋg])<sub>Stem</sub> al      ('diphtho[ŋ])<sub>LWd</sub>

In addition, stress-placement patterns may differ for Level I affixed stems in comparison with their corresponding independent words. In (3), the main stress of the LWd occurs on the initial syllable, but on the final syllable when the stem is followed by the Level I suffix.

As we discuss in §2.1.2, differences in the phonological behaviour of stems can be accounted for by postulating word-level boundaries before Level II suffixes but not before Level I suffixes. However, there are also other ways to account for these differences, as we will see in §2.1.3. In §§2.4 and 2.5 we discuss experiments designed to test the word boundary proposal.

<sup>1</sup> Other examples (but not a complete list) of Level I suffixes are *-ic* (e.g. *'atom vs. d'tomic*), *-ity* (e.g. *'curious vs. curi'osity*), *-ation* (e.g. *con'de[m] vs. .conde[m.n]ation*).

(4) a. MWd  
 |  
 Root  
 |  
 Root    Level I suffix

b. MWd  
 /    \  
 MWd    Level II suffix  
 |  
 Root

2.1.3 *The PWD boundary proposal.* Following the Indirect Reference Hypothesis that phonological rules only refer to phonological structure but not directly to morphosyntactic structure (Selkirk 1986, Inkelas 1990), Selkirk's (1984) MWd *vs.* Root distinction for stems with Level I *vs.* Level II suffixes can be translated into representational differences in prosodic constituent structure.<sup>2</sup> MWd boundaries that coincide with the

<sup>2</sup> There is a great deal of evidence from a variety of languages to support this hypothesis. For example, the domain of Mandarin Chinese tone sandhi is related, but not necessarily isomorphic, to morphosyntactic constituent structure (Chen 1990).

edges of Level II stems should have prosodic counterparts, i.e. the edges of some prosodic domain  $\alpha$ , from which Level II stems are excluded. In contrast, Root-level stems and their Level I suffixes coexist within the same prosodic domain  $\alpha$ . Aronoff & Sridhar (1983) and Szpyra (1989), among others, propose that this prosodic domain  $\alpha$  is a PWd.

A word-level phonological (prosodic) constituent was originally proposed by Selkirk (1980a) to account for main word stress assignment in English. According to her, binary branching prosodic constituent structure above the syllable, i.e. the foot and the prosodic word, determines weak-strong relations among syllables. On this view, the location of main stress is ultimately derived from rules relating to these constituents: in English, the leftmost syllable in a foot is strong, i.e. stressed, and a strong syllable dominated by the strongest foot of a PWd bears the main word stress. The fact that Level II suffixes do not affect the placement of main word stress on stems is due to their stems forming an independent PWd.

The PWd has also been proposed to be the domain for segmental rule application (Selkirk 1980a, b, Nespor & Vogel 1986). There are even rules that apply only to PWd junctures, such as voicing assimilation and deaspiration in Sanskrit (Selkirk 1980b). The consonant-cluster simplification observed in English Level II suffixation cases is also captured as a PWd juncture phenomenon.

The adjunction of Level II stems to a PWd-level constituent can be described in Optimality Theory terms by a highly ranked constraint requiring a string of segments that correspond to Level II suffixes to be immediately preceded by a PWd base, introduced by McCarthy & Prince (1993b: 67ff) as SUFFIX-TO-PWD, and modified as in (5).<sup>3</sup>

#### (5) LEVELII-SUFFIX-TO-PWD

The terminal string of segments in the output representation, corresponding to a Level II suffix in the input, is immediately preceded by a PWd. That is, the base of Level II suffixation is a PWd.

If LEVELII-SUFFIX-TO-PWD is undominated in English, Level II suffixes should be preceded by a PWd boundary regardless of their segmental content and organisation. However, Raffelsiefen (2005: 234) claims that

---

Ferreira (1993) shows that word and pause durations in English are not successfully predicted by syntactic structure alone: they are better explained if prosodic structure is posited.

<sup>3</sup> We do not adopt here a morphology-prosody right-edge alignment constraint, ALIGN-R(MWd, PWd) or ALIGN-R(Level II Stem, PWd), which calls for alignment between the right edge of every MWd stem (or Level II stem) and the right edge of a PWd. This is because the constraint would be violated by surface representations that undergo stem-final consonant-cluster simplification, e.g. ((*condem*)[m])<sub>PWd</sub> *ing*). Since the rightmost consonant /n/ of the stem *condemn* is absent in the surface representation of the stem (*condem*[m])<sub>PWd</sub>, the PWd boundary is misaligned to the stem-medial consonant /m/, resulting in the violation of the ALIGN constraint above. LEVELII-SUFFIX-TO-PWD, however, is not violated by the representation with stem-final cluster simplification, because the constraint does not refer to the segmental content of the PWd preceding the suffix.

stems with a single obstruent suffix, e.g. *girl-s*, should have no PWd break between the right edge of the stem and the suffix, as shown in (6a). According to her, the representation in (6b), which satisfies LEVELII-SUFFIX-TO-PWD, should be ruled out, on the assumption that all segments must be parsed into syllables, i.e.  $\text{PARSE}(\text{seg})$  ('every segment is parsed into a syllable') is undominated.

- (6) a.  $((\text{girls})_{\sigma})_{\text{PWd}}$       b.  $*((\text{girl})_{\sigma})_{\text{PWd}} \text{ s}$

However, we do not adopt Raffelsiefen's view, because the assumption that  $\text{PARSE}(\text{seg})$  is undominated is a stipulation without empirical support. Rather, there is evidence that may support the view that the suffix *-s* in English is actually extrasyllabic. Walsh & Parker (1983) found that the suffix *-s*, as in (6b), was longer than word-final *s* in monomorphemic forms. On the assumption that segment duration relates to prosodic structure, their observation suggests that the suffix *-s* forms part of a different prosodic structure than corresponding segments in monomorphemic forms. It therefore appears possible to represent the suffix *-s* as phonologically extrasyllabic.

Hall (2001) and Raffelsiefen (2005) further claim that there is no PWd boundary between a stem and a vowel-initial Level II suffix such as *-ing*, and propose that a PWd boundary should be present only before Level II suffixes such as *-ness* and *-less*, which begin with a consonant. However, as noted in §2.1.1, Level II suffixes that begin with vowels induce similar phonological behaviour to consonant-initial Level II suffixes, i.e. final consonant cluster simplification, as shown again in (7a), and stress neutrality, as in (7b).

- (7) a.  $(\text{si}/\eta\text{g}/ \rightarrow [\eta])_{\text{Stem}} \text{-ing}_{\text{V-Ons}}$        $(\text{so}/\eta\text{g}/ \rightarrow [\eta])_{\text{Stem}} \text{-less}_{\text{C-Ons}}$   
       b.  $(\text{'parent})_{\text{Stem}} \text{-ing}_{\text{V-Ons}}$        $(\text{'parent})_{\text{Stem}} \text{-less}_{\text{C-Ons}}$

These phenomena provided the original motivation for a word-level boundary before Level II suffixes, regardless of whether they are vowel- or consonant-initial, *contra* Raffelsiefen and Hall.

It is nevertheless worthwhile considering why Raffelsiefen and Hall argue against a PWd boundary before vowel-initial suffixes. According to Hall (2001), these suffixes absorb the final consonant of the immediately preceding stem as their syllable onset, as shown in (8) for *sunning*, and this configuration is claimed to be incompatible with a PWd boundary between the stem-final consonant and a Level II suffix.

- (8)  $(\text{s}\Lambda)_{\sigma 1} (\text{n}\eta\text{n})_{\sigma 2}$   
       ↑  
       stem-final consonant  
       = onset of  $\sigma 2$

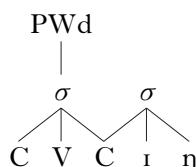
However, flapping phenomena in American English suggest that stem-final consonants immediately preceding these vowel-initial Level II



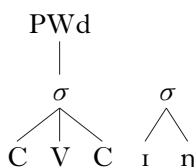
suffixes are not completely integrated into the syllable of the suffix. For example, the stem-final consonant /t/ in *date* surfaces as a flap when followed by *-ing* and *-er*, as in *da[r]ing* and *da[r]er*, and is therefore very different from the realisation of syllable-initial /t/ in word-initial and pretonic positions. The flapping of an intervocalic alveolar stop between a stressed vowel and an immediately following unstressed syllable nucleus has been proposed to be due to its being ambisyllabic, i.e. being simultaneously the coda of the preceding syllable and the onset of the following syllable (Kahn 1976). Selkirk (1982b), however, claims that it is unambiguously syllabified into the preceding stressed syllable. Turk (1993, 1994) presents experimental evidence that suggests that intervocalic consonants in this context are more similar in their articulatory characteristics to syllable-final than to syllable-initial stops. Turk (1994), in particular, lends support to Selkirk's (1982b) view. She compared the upper lip gestures of clearly syllable-initial labial stops (e.g. *re'pair*), those of clearly syllable-final stops (e.g. *'captor*) and those of stops immediately preceded by stressed vowel and followed by unstressed nucleus (e.g. *'leper*), whose syllabic affiliation is not known. She found that the upper lip gestures of the unknown stops patterned with those of the clearly syllable-final ones.<sup>4</sup>

Contrary to Hall's (2001) proposal in (8), Kahn's (1976) and Selkirk's (1982b) views of the syllabification of post-tonic intervocalic consonants into the preceding tonic syllable allow for a PWD boundary between the stem and the suffix, as shown in (9).

(9) a. ambisyllabicity  
(Kahn 1976)



b. unambiguous syllabification  
(Selkirk 1982b)



In the representation in (9a), the ambisyllabic affiliation of the stem-final consonant violates the CRISPEDGE( $\sigma$ ) constraint in (10), which prohibits double linking of segments at syllable edges (Itô & Mester 1994).

<sup>4</sup> The kinematic measurements of the tongue tip, upper lip and lower lip in Turk (1993) were consistent with the upper lip analyses in Turk (1994), in that the consonant articulations whose syllable affiliation were ambiguous were clearly more coda-like than onset-like. However, there were some indications in the data of differences between 'unknown' consonant articulations and those of clear codas. These differences varied across articulators and in type: sometimes unknown consonants formed a clear third category, sometimes some patterned like onsets and others like codas. Although difficult to interpret, these patterns might argue in favour of an ambisyllabic analysis.

(10) CRISPE<sub>EDGE</sub>( $\sigma$ )

Segments that coincide with the edges of  $\sigma_i$  must be exhaustively contained in  $\sigma_i$ .

Nonetheless, the representation in (9a) may still surface as optimal, provided that ONSET, i.e. a constraint that calls for an onset consonant in every syllable, dominates CRISPE<sub>EDGE</sub>( $\sigma$ ), as suggested by Itô & Mester (1994), who claim that the flapping of word-final /t/ immediately followed by a word-initial vowel, as in *sough*[ɾ] *Ed*, is due to its being ambisyllabic.

We are not in a position to choose between the representations in (9a) and (9b). The significant point here is that both of these approaches allow for a PWd boundary between the stem and the following Level II suffix. One of the goals of the current study is to test the possibility that a PWd boundary occurs between a stem and a Level II suffix, even when the suffixes are monoconsonantal (e.g. -s) or begin with a vowel (e.g. -ing).

**2.1.4 Approaches without a PWd boundary: paradigm uniformity and serialism.** In Raffelsiefen's (2005) approach, where a PWd boundary does not occur before unstressed vowel-initial Level II suffixes, the coda-cluster simplification observed in e.g. *conde*[m]ing should be due to Paradigm Uniformity or output-output (OO) correspondence.

The principal tenet of the OO-correspondence approach is that the phonological similarity between two related forms in the same paradigm (e.g. a simplex word and a suffixed word) comes from a mapping relationship between the output forms of suffixed words and those of their base (e.g. simplex words). According to Benua (1997), 'word-boundary' phenomena, e.g. stem-final consonant-cluster simplification before Level II suffixes, are formalised by adopting 'OO-Identity' constraints which require the optimal output of the stem in the affixed word (e.g. *condemn* in *condemning*) to be identical to the surface representation of its base form (the stem forming an independent word on its own without affixes, e.g. (*conde*[m])<sub>PWd</sub>, accompanied by word-final consonant-cluster simplification).

In Benua's model, there are two kinds of OO-Identity constraints: one referring to Level I affixation and the other referring to Level II affixation. Level I OO-Identity constraints and Level II OO-Identity constraints are ranked differently with respect to other types of constraints, such as input-output (IO) constraints, i.e. constraints requiring the output string to have identical representations to its input. Since Level I OO-Identity constraints are dominated by IO constraints, the surface representation of the Level I suffixed stem retains all the segments from its input representation, e.g. *conde*[mn]ation. In contrast, Level II OO-Identity constraints outrank the IO constraints, and therefore result in consonant-cluster simplification at the end of the Level II suffixed stem. Similarly, the highly ranked Level II OO-Identity constraints force the stress location of Level II suffix forms and base forms to be identical. In this

way, the OO-Identity approach captures the difference in the behaviour of Level II and Level I suffixed stems, without resorting to postulating differences in their morphological and phonological constituent structure.

Another approach that does not require a PWd boundary before Level II suffixes is a serial derivational approach (e.g. Siegel 1974, Kiparsky 1982, Halle & Mohanan 1985, Borowsky 1993).<sup>5</sup> According to Borowsky (1993), for example, Level I suffixes attach to their stems earlier in the derivation, and the combination of stem + Level I suffix undergoes word-level phonological operations as a single unit, while Level II suffixes attach to their stems after the stems have independently undergone word-level phonological operations such as stress assignment and final coda-cluster simplification. It has also been proposed that internal constituent boundaries are erased after completion of phonological rule application at each lexical level, i.e. 'bracket erasure' (Kiparsky 1982: 140), so that phonological processes at later levels are inaccessible to any constituent structure constructed during earlier levels of derivation. In this approach, the phonological differences between Level I and Level II suffixed forms such as *condemn-ing* and *condemn-ation* are explained by ordered rules that apply to different morphological levels, rather than by the necessary occurrence of a PWd boundary before the Level II suffix.

In summary, there is no clear agreement as to the existence of word-internal PWd boundaries in Level II suffixed forms. The main goal of the experiments presented in what follows is to test for durational evidence for such boundaries. Segment duration is investigated, since this is known to vary systematically according to the presence of a prosodic boundary, and its level within the prosodic hierarchy (e.g. Wightman *et al.* 1992). Experiment 1 tests for durational differences between monomorphemic words and Level II suffixed forms, and Experiment 2 tests for differences between Level I suffixed forms and Level II suffixed forms. Word-internal PWd boundary proposals predict that there should be durational differences in both cases.

## **2.2 Testing the presence *vs.* absence of a PWd boundary before Level II suffixes**

Our experimental strategy was to compare the rhyme durations of stems followed by Level II suffixes (e.g. /ak/ in *tack-s*, /ek/ in *bake-ing*) with those of corresponding stretches in monomorphemic words (e.g. /ak/ in *tax*, /ek/ in *bacon*). We chose this strategy based on our assumption that prosodic constituent structure has measurable durational consequences for phonetic segments in processes such as polysyllabic shortening, poly-segmental shortening and pre-boundary lengthening. At the phrase level (i.e. above the word level), pre-boundary lengthening in particular has

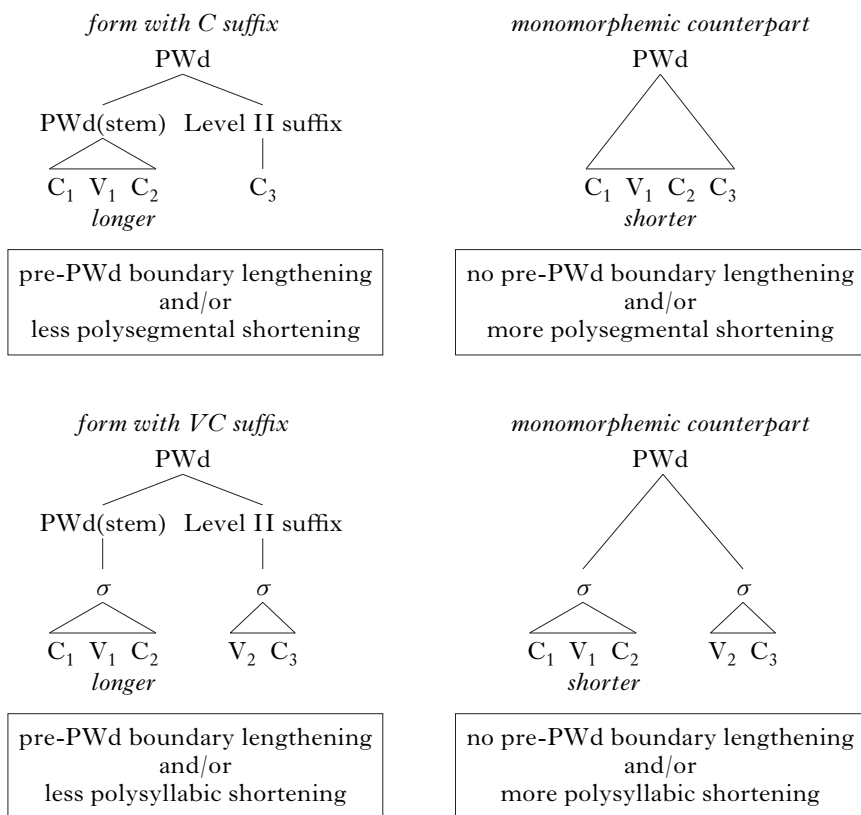
<sup>5</sup> Kiparsky (2000, 2008) and Bermúdez-Otero (1999) reinterpret the rule-based serialist approach in the framework of constraint-based phonology, i.e. stratal OT, in which stem level and word level are associated with separate constraint systems.

been well documented (e.g. Wightman *et al.* 1992, Cho & Keating 2001). Word-level durational adjustments are also well described (e.g. Lehiste 1972, Huggins 1975, Beckman & Edwards 1990, Turk & Shattuck-Hufnagel 2000), and can be accounted for by a variety of processes, including polysyllabic shortening and perhaps pre-word boundary lengthening, at least in accented contexts (Turk & Shattuck-Hufnagel 2000). For example, Turk & Shattuck-Hufnagel observed that the rhyme duration of a monosyllabic word, e.g. the duration of /un/ in *tune acquire*, was about 30 ms longer than that of the corresponding rhyme of a disyllabic word, e.g. the duration of /un/ in *tuna choir*, when the syllable containing the rhyme was pitch-accented. This can be interpreted as being due to the polysyllabic shortening of /un/ in the disyllabic word, and/or to PWd-final lengthening in the rhyme of the monosyllabic word.

It is beyond the scope of this paper to determine the mechanism(s) (e.g. polysyllabic shortening or pre-boundary lengthening) responsible for observed duration differences (cf. Turk & Shattuck-Hufnagel 2000 for a discussion). Rather, what is crucial here is that the structural difference between *tune acquire* and *tuna choir* predicts the durational differences observed on the target segments.

Several other types of durational adjustments are also proposed to be related to prosodic structure. One of these, polysegmental shortening, accounts for segments being shorter when there are more consonants in a syllable (Campbell & Isard 1991). For example, Lehiste (1960) and Christie (1977) observed shorter consonant durations in English when the consonants belonged to a coda cluster as compared to a heterosyllabic cluster, e.g. /m/ was longer in *plum pie* than in *plump pie*. Waals' (1999) study of Dutch showed that a vowel followed by a singleton coda in a monosyllabic word was 10 ms longer than one followed by a coda cluster. She also observed that Dutch /k/ was 16 ms longer when it was the only consonant in the coda of a monosyllabic word, as in CVk, than when it was part of a coda cluster, as in CVks. Note that this example can be accounted for either by pre-boundary lengthening (longer /k/ before a word boundary than before /s/) or by polysegmental shortening (shorter /k/ in a cluster than in a singleton). Here again, the two duration-implementation mechanisms make similar predictions (see similar discussion in Turk & Shattuck-Hufnagel 2000).

The crucial point for our purposes is that all of these mechanisms make the same predictions for stimuli proposed to contain word-internal PWd boundaries, that is, longer stem-rhyme durations for Level II suffixed forms than for corresponding stretches of monomorphemic forms. That is, we expect PWd-final lengthening, or less polysegmental/polysyllabic shortening of the stem syllable, as compared to the corresponding stretch in a monomorphemic word of the same number of syllables. The predictions for polysegmental and polysyllabic shortening are based on the assumption that the magnitude of polysyllabic and polysegmental shortening correlates with the hierarchical organisation of prosodic constituent structure, as is the case for pre-boundary lengthening. Wightman *et al.*



*Figure 1*

Prosodic constituent structures with PWd boundaries before Level II suffixes,  
and predictions by duration-implementation mechanisms.

(1992) show that prosodic constituent boundary depth is proportional to the degree of pre-boundary lengthening: the higher the following prosodic boundary, the greater the pre-boundary lengthening. By analogy, we assume that the magnitude of polysyllabic and polysegmental shortening is attenuated by prosodic boundaries between syllables or segments, and that the magnitude of attenuation correlates with the strength of the boundary: the stronger/higher the prosodic boundary between them, the less polysyllabic/segmental shortening operates on those syllables and segments. If there is a PWd boundary between the syllable/segments of the stem and those of the Level II suffix, the effect of possible polysyllabic and polysegmental shortening of those syllables/segments should be attenuated (see Fig. 1).<sup>6</sup>

<sup>6</sup> We adopt the assumption in Fig. 1 that forms with Level II suffixation have a recursive PWd structure in which PWd(stem) together with the following suffix



Figure 2

Prosodic constituent structure: (a) nested PWd structure;  
 (b) sequence of two PWds.

From the proposed duration-adjustment mechanisms and associated assumptions, we predict that the duration of a stem followed by a Level II suffix should be longer than the corresponding sequence in a monomorphemic counterpart, all other things being equal. So for example, we expect the rhyme duration in stems (e.g. /ak/ in *tack-s*, /ek/ in *bake-ing*) to be greater than the duration of the corresponding stretches in monomorphemic forms (e.g. *tax*, *bacon*).

Note also that the nested status of the proposed PWd structure (see note 6) leads us to predict smaller phonetic duration differences for our comparisons than for polysyllabic shortening effects described in the literature for e.g. *tune acquire* vs. *tuna choir* (about 30 ms: Turk & Shattuck-Hufnagel 2000) and polysegmental shortening effects (10 ~ 25 ms: Waals 1999). As shown in Fig. 2, only one PWd boundary appears after the stem in the nested PWd structure (Fig. 2a), whereas two PWd boundaries follow a monomorphemic word when followed by another content word (Fig. 2b). In previous studies of phonetic duration differences due to word boundary structure (e.g. Turk & Shattuck-Hufnagel 2000), comparisons were made between (i) cases where there are *two* word boundaries immediately after the target rhyme or syllable, e.g. /tun/<sub>PWd</sub> /əkwaɪɪ/<sub>PWd</sub> *tune acquire* (i.e. the right edge of the preceding word *tune* and the left edge of the following word), and (ii) cases without a boundary immediately after the target rhyme or syllable, e.g. /tunə/<sub>PWd</sub> /kwaɪɪ/<sub>PWd</sub> *tuna choir*.

It is likely that the magnitude of any durational adjustments to the target rhyme will be positively correlated with the number of prosodic word boundaries immediately after the target syllable or segments. As a result, we expect less pre-word boundary lengthening and/or less attenuation of polysyllabic and polysegmental shortening for the nested PWd structure shown in Fig. 2a (tested here) than in the two-word sequence in Fig. 2b (reported in the literature).

---

syllable are dominated by another PWd. Such recursive structure violates one of the constraints on prosodic hierarchy, NON-RECURSIVITY (Selkirk 1995), i.e. no prosodic constituent  $C_i$  is dominated by another prosodic constituent  $C_j$  when  $C_i$  and  $C_j$  belong to the same level in the prosodic hierarchy. NON-RECURSIVITY is a violable constraint, as discussed by Ladd (1986), McCarthy & Prince (1993a, b) and Selkirk (1995), among others.

## 2.3 Relevant findings from previous studies

Previous durational studies of stems and a Level II suffix *-s* are consistent with the word-internal PWd-boundary hypothesis.

Walsh & Parker (1983) found that the Level II suffix *-s* [s] in three suffixed words (*wreck-s*, *lap-s* and *heart-s*) was 10 to 12% (about 9 ms) longer than its monomorphemic counterpart in *Rex*, *lapse* and *Hartz*. Schwarzlose & Bradlow (2001) also found that stem-final consonants in four suffixed words ([k] in *tack-s*, *tuck-s* and *mac-s* and [t] in *hurt-s*) were 2 to 6% (about 3 to 5 ms) longer than the penultimate segments [k] in *tax*, *tux* and *max* and [t] in *Hertz*.

Sproat (1993) and Sproat & Fujimura (1993) studied the influence of Level I *-ic* vs. Level II *-ing* on the duration of segments in the stem *beel*. They revealed no significant durational differences, although the direction of the effect was as predicted by the PWd-boundary hypothesis, that is, longer [il] in *beeling* than in *beelic*.<sup>7</sup>

While generally consistent with the PWd-boundary hypothesis, effects reported in previous studies were small, were based on a small number of tested items (at most four in Schwarzlose & Bradlow 2001) and were not always statistically reliable. In addition, studies of forms with *-s* by Walsh & Parker and Schwarzlose & Bradlow were based on comparisons of homophones, where subjects could have artificially introduced a contrast between members of homophonous pairs.

Finally, four out of seven items with the suffix *-s* used by Walsh & Parker and Schwarzlose & Bradlow had possible spelling confounds. Pairs such as *max* vs. *macs*, *tux* vs. *tucks* and *Rex* vs. *wrecks* differ in the number of letters they contain. In particular, the target coda consonants in monomorphemic forms correspond to only part of a single letter, whereas the corresponding coda consonants in suffixed forms correspond to whole letters. Treiman & Cassar (1997) found that orthography has an effect on perceived phoneme count, e.g. a sequence of two phonemes spelled with two letters was judged to contain more 'sounds' than that spelled with one letter, and Warner *et al.* (2004) found that acoustic closure durations (i.e. a silent period due to the closure) of Dutch (non-geminate) consonants spelled with two letters, e.g. /d/ in *baadden* 'they bathed', were 3 ms longer than consonants spelled with a single letter, e.g. /d/ in *baden* 'to bathe'.

## 2.4 Experiment 1

The purpose of this experiment is to test the prediction of the word-internal PWd-boundary hypothesis that stem-rhyme duration in suffixed forms is longer than the corresponding stretch in monomorphemic forms. This study is designed to test the generality of Schwarzlose & Bradlow's

<sup>7</sup> As noted by the associate editor, since *beelic* (*beel-ic*) is a nonce word, pseudo-productively created by Sproat & Fujimura (1993), the stem *beel-* in *beelic* may have been forced into the Level II category.

greed/greet/niece	i	put/broad/brute/Bruce	ʊ
agreed/sea/freeze	i:	brewed/bruise/moo	u:
grade/greyed/great/say	e ↔ e:	queue	jʊ:
bid	ɪ ↔ ə	but	ʌ
bed	ɛ	side/price	Δi
Pam	a ↔ ɑ	sighed/prize	ɑ:e
cot/caught	ɔ	bough	ʌʊ
coat/code/flowed/sew	o ↔ ɔ:	boy	ɔe

Table I

The phonemes and major allophones of Scottish Standard English, adapted from Scobbie *et al.* (2006: 7). The arrows indicate ‘a particularly notable range of productions’.

(2001) findings, using a greater range of items, i.e. not only forms with *-s* [s] (the plural or 3rd person singular suffix that attaches to voiceless non-sibilant consonant-final nouns or verbs), but also with *-ing* (the gerund suffix), *-ed* [t] (the past tense suffix that attaches to verbs ending with non-alveolar voiceless consonants) and *-es* [ɪz] (the plural or 3rd person singular suffix that attaches to sibilant-final nouns or verbs). While homophonous pairs such as *mix vs. micks* were included in our dataset to make them parallel to those in the previous literature, it also included test sequence comparisons spelled with identical numbers of letters. Factors such as pitch-accent location and speech rate were controlled throughout.

**2.4.1 Speakers.** Five native female speakers of Scottish Standard English (SSE) from Edinburgh served as paid subjects in the experiment. None reported any hearing or speaking difficulties. They were all University of Edinburgh undergraduate students (aged between 18 and 40) at the time of the recording.

The phonemic inventory of SSE provided in Table I. The most notable durational features of SSE are the patterns arising from the Scottish Vowel Length Rule (SVLR), which applies primarily to the vowels /i ʊ Δi/ (Scobbie *et al.* 1999, Scobbie *et al.* 2006), lengthening them immediately before lexical word boundaries, Level II morpheme boundaries, voiced fricatives and /r/. They are short elsewhere. Therefore, both *brute* and *brood* are pronounced with a short vowel [ʊ], while *brewed* (*brew#ed*) has long [u:]. According to Scobbie *et al.* (1999), the difference in duration between vowels in the short and long contexts is 40 to 80 ms.

In our experiment, we avoid words that would undergo the SVLR. Potential durational differences between Level II suffixed stems and their monomorphemic counterparts in our experiments are therefore expected to be much more subtle than those of the SVLR.



Set 1	stem + suffix -s [s]	monomorph	Set 2	stem + suffix -ed [t]	monomorph
[t <sup>h</sup> aks]	tacks	tax	[mɪst]	missed	mist
[t <sup>h</sup> ʌks]	tucks	tux	[p <sup>h</sup> ast]	passed	past
[laps]	laps	lapse	[p <sup>h</sup> est]	paced	paste

Set 3			Set 4		
stem + suffix <i>-ing</i> [ɪŋ]		monomorph	stem + suffix <i>-es</i> [ɪz]		monomorph
baking	[bekɪŋ]	bacon	[bekɪ]	axes	[aksɪz]
canning	[kʰanɪŋ]	cannon	[kʰanɪ]	axis	[aksɪs]
raising	[ɹeɪzɪŋ]	raisin	[ɹeɪzɪ]	bushes	[bʊʃɪz]
summing	[sʌmɪŋ]	summit	[sʌmɪt]	bushel	[bʊʃl]
socking	[sɔkɪŋ]	socket	[sɔkɪt]		
puffing	[pʰʌfɪŋ]	puffin	[pʰʌfɪn]		
bobbing	[bɔbɪŋ]	bobbin	[bɔbɪn]		

*Table II*

Monosyllabic (Sets 1 and 2) and disyllabic (Sets 3 and 4)  
item pairs used in Experiment 1.

**2.4.2 Stimulus design.** The stimulus sets consisted of (i) bimorphemic items consisting of closed syllable stems followed by a Level II suffix and (ii) their monomorphemic counterparts. Level II suffixes included the voiceless plural or 3rd person singular suffix -s [s], the voiceless past tense suffix -ed [t] and two syllabic suffixes: the gerundive suffix -ing [ɪŋ] and the plural inflectional suffix -es [ɪz]. Test pairs therefore included monosyllables such as *tack-s* vs. *tax* and *miss-ed* vs. *mist*, as well as disyllables such as *bake-ing* vs. *bacon*, *puff-ing* vs. *puffin* and *ax-es* vs. *axis*.

We emphasise here that the test sequences in disyllabic pairs crucially contain the same number of letters, and are therefore not confounded by potential spelling effects on duration: pairs like *puff-ing* and *puffin* have the same number of letters for their target VC sequences. Non-homophonous pairs like *baking* vs. *bacon* were included to ensure that any durational differences between the two morphological conditions would not be due to speakers artificially contrasting members of homophonous pairs within the experimental task.

Three pairs were chosen for each of the monomorphemic word comparison sets, e.g. *tack-s* vs. *tax* (Set 1) and *miss-ed* vs. *mist* (Set 2), seven pairs for one of the disyllabic word comparison sets, e.g. *bake-ing* vs. *bacon* (Set 3), and two pairs for the other disyllabic word comparison set, e.g. *ax-es* vs. *axis* (Set 4). The four sets of test pairs (15 pairs in total) are shown in Table II. No proper names were included in the stimulus sets, because an earlier pilot study suggested that proper names tend to be longer than words spelled without capitals.

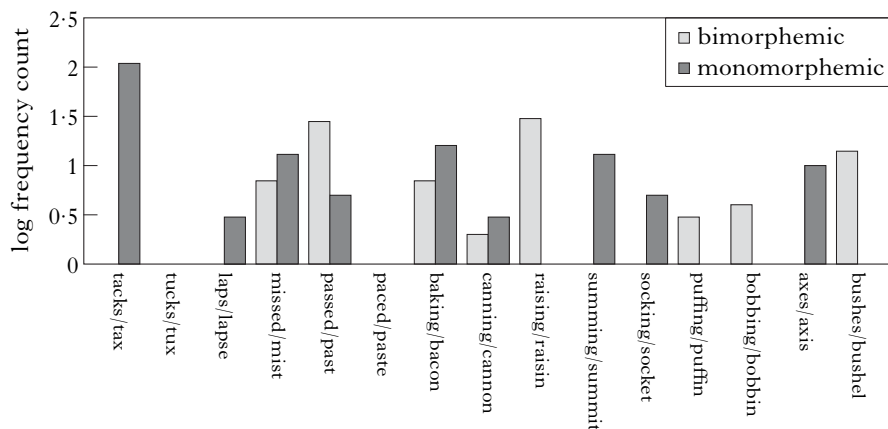


Figure 3

Log word frequency count of items used in Experiment 1.

We provide in Fig. 3 the log word frequencies of the stimulus pairs, which are obtained from the CELEX database (Baayen *et al.* 1993). Lexical frequency is known to contribute to segment duration (e.g. Munson 2001, Bell *et al.* 2002, Aylett & Turk 2004, Pluymaekers *et al.* 2005). For example, Aylett & Turk (2004) show that words whose log frequency exceeds 4 to 4.5 are realised with shorter durations than less frequent ones. Since we limited the log frequency count of each item to below 2.5, we expected little effect of frequency on the durations of our items.

**2.4.3 Elicitation and recording.** Each set of sentences was printed on a set of cards and presented to speakers in random order (by shuffling). Our strategy was to elicit target words in contexts where the subtle durational effects expected here would be most likely to occur. Because the durational effects of word-constituent structure are known to be larger in pitch-accented contexts than non-pitch-accented contexts (Beckman & Edwards 1990, Turk & Shattuck-Hufnagel 2000), background and carrier sentences were designed to elicit nuclear pitch accents on target words (by contrastive focus for Sets 1 and 2 and presentational focus for Sets 3 and 4), as illustrated in the examples in (11) (the full set of test materials is given in the Appendix).

- (11) a. A special event at the cooking institute. Say '*canning* event' for me.  
 b. The 1 o'clock event at the Edinburgh Castle. Say '*cannon* event' for me.

Target words are italicised in (11), but were not highlighted in any way when presented to subjects. The topic sentence in (11) states that the

following carrier sentence refers to a particular event, and the target word in the carrier sentence further specifies the type of event. The target word is therefore interpreted as new (i.e. in presentational focus), and speakers put a nuclear pitch accent on the target word, while the words following the target word, i.e. *event for me*, are unaccented. Within each carrier sentence, the target word was immediately preceded by *Type* (when the target word began with a vowel or a rhotic) or *Say* (when the target word began with another consonant). This first word in the carrier sentence will henceforth be called the head. The target word was also followed by one or more words (the post-target). The head and the post-target were kept constant between the two utterances in each comparison pair.

We elicited our materials at different rates, on the assumption that boundary effects would be magnified at slow speech rates. Beckman & Edwards (1990) found that the duration difference of the [a] in *pop opposed* vs. *poppa posed* was magnified by about 50 ~ 100 % at a slower speech rate.

Our recording consisted of multiple blocks, where each block was read at either a normal or a slow rate. Only the blocks consisting of pairs from Sets 3 and 4 were read at an extra-slow rate in addition to the normal and slow speech rates. When reading at the normal rate, speakers were instructed not to put any pause between words in a single sentence; none of the speakers had difficulty in doing this, i.e. they read the sentences fluently. If they stumbled over the pronunciation of a sentence, they were instructed to read the whole sentence again from the beginning, and the disfluent speech was discarded. When reading at the slow rate, speakers were instructed to read not only slowly but also articulately ('please say these phrases as if you are talking to someone having difficulty in hearing or to small children'), and for the extra-slow rate they were told to read even more slowly than for the slow rate speech produced in the previous block. The use of articulate speech helped speakers to slow down their speech rate. At the slow rates, speakers were allowed to put pauses between words.

The two stimuli in each comparison pair (e.g. the sentences in (11a) and (b)) always occurred together in the same block, but were never presented to speakers consecutively. For each of the speech rates, six repetitions were elicited for stimulus pairs in Sets 1 and 2 for four of the five speakers; four repetitions were elicited in all other cases. Note that not all pairs were read by all of the five speakers. Two of the nine pairs in Set 3 (i.e. *puffing* vs. *puffin* and *bobbing* vs. *bobbin*) and the two pairs in Set 4 (i.e. *axes* vs. *axis* and *bushes* vs. *bushel*) were read by four of the five speakers.

Recordings were made in a sound-treated recording studio at the Department of Theoretical and Applied Linguistics, University of Edinburgh, using a hypercardioid microphone and a SONY PCM-2700A DAT recorder. Recordings were amplified, low-pass filtered at 48 kHz and downsampled to 16 kHz for further analysis.

**2.4.4 Measurements and analyses.** We measured the rhyme duration of the stem of each target word, as well as the corresponding sequence in the

monomorphemic counterpart. For example, the target sequence for the monosyllabic words *tacks* and *tax* [t<sup>h</sup>aks] was the sequence [ak], i.e. the nucleus vowel and the final coda consonant of the stem. In the same way, the target sequence for *baking* and *bacon* (i.e. [bekɪŋ] and [bekn̩]) was the sequence [ek].

We defined the target sequence acoustically as starting at the onset of the word-initial consonant-constriction release and ending at the onset of stem-final coda consonant-constriction release (or the onset of the corresponding coda consonant in the monomorphemic counterparts).<sup>8</sup> Therefore, the period of aspiration immediately following the constriction release of a word-initial voiceless stop was also included in the target sequence. We chose this criterion because the release of each word-initial consonant can be interpreted as the onset of supralaryngeal articulation of the following nucleus vowel (see Turk *et al.* 2006 for discussion).

Our speakers pronounced the initial rhotic of the *raisin* and *raising* pair as an approximant [ɹ]. The segmentation of prevocalic approximant variants is not straightforward, because voicing continues from the approximant to the following vowel, formant trajectories are continuous throughout the approximant–vowel sequence and amplitude does not change abruptly from the approximant to the vowel. [ɹ] was therefore included in the target-measurement sequence.

Part of the initial word, i.e. the head, of each carrier sentence (*Say* or *Type*) was also measured, as a check for overall speech-rate differences between test conditions. The interval measured for *Say* began at the frication offset of the onset /s/ and ended at the F2 offset of the vowel. For *Type* the interval began at the onset of constriction release for /t/ and ended at the onset of constriction release for /p/. These control intervals were therefore comparable to the test intervals in size.<sup>9</sup>

Because each test sentence in the suffixed condition was paired with a corresponding test sequence in the monomorphemic condition in each recording block, we analysed our data using paired t-tests. One-tailed tests were adopted instead of two-tailed tests, because we were testing the hypothesis that the duration of one condition is longer than the other (rather than assessing a more general hypothesis of a difference in either direction).

**2.4.5 Results 1: speech-rate index.** The mean duration of the control segment (i.e. the rhyme of the initial word of each carrier sentence: *Say*

<sup>8</sup> Strictly speaking, ‘constriction release’ is an articulatory term. Nonetheless, we use the term as an acoustic measurement index because oral constriction and its release are reflected acoustically. For example, stop consonant constriction release corresponds to noise burst acoustically. Fricative consonant constriction release corresponds to the end of frication energy, and the release of a lateral consonant /l/ is acoustically marked as an abrupt increase in amplitude.

<sup>9</sup> The best way to assess differences in speech rate across utterances within experimental blocks is a matter of debate. The method that we adopt here has been used in the literature, e.g. by Turk & Shattuck-Hufnagel (2000).

and *Type*) of the stem-suffix cases and that of the monomorphemic cases were compared, using a two-tailed paired t-test, for each comparison pair at each speech rate. Results suggested that speech rate was comparable across experimental conditions, i.e. the control-segment duration mean of the stem-suffix cases was neither significantly longer nor significantly shorter than that of the monomorphemic cases.

The only exception was Set 4 at the normal speech rate: the control-segment duration mean for the stem-suffix cases was 2.3% (i.e. 5 ms, SD = 18) longer than that of the monomorphemic case, which was marginally significant in a two-tailed paired t-test ( $\alpha = 0.05$ ): (control-segment mean for the stem-suffix cases = 174 ms (SD = 63), control-segment mean for the monomorphemic cases = 169 ms (SD = 60);  $t = 1.68$ ,  $df = 47$ ,  $p = 0.1$ ).

We therefore consider it unlikely that any observed systematic differences between our test sequences across experimental conditions are due to differences in overall speech rate, with the possible exception of the differences found in Set 4 (at the normal rate only).

**2.4.6 Results 2: test sequences.** The differences observed between stem-suffix and monomorphemic cases were in the direction predicted by the word-internal PWD-boundary hypothesis for some sets at the normal rate, and for all sets at the slow speech rate.

*The normal rate.* The differences observed at the normal speech rate were in the direction predicted by the PWD hypothesis, but the magnitudes of differences were very subtle, and reached statistical significance only for Set 1. For Set 1 (e.g. *tacks* vs. *tax*), the initial syllable rhyme duration of the suffixed words (mean = 218 ms, SD = 34) was on average 1.9% (4 ms, SD = 16) longer than the test sequence in monomorphemic words (mean = 214 ms, SD = 32;  $t = 2.331$ ,  $df = 82$ ,  $p = 0.011$ ). For Set 2 (e.g. *missed* vs. *mist*), the rhyme duration of the suffixed words (mean = 232, SD = 54) were again on average 1.8% (4 ms, SD = 23) longer than that of the monomorphemic words (mean = 228, SD = 51;  $t = 1.60$ ,  $df = 83$ ,  $p = 0.057$ ), where the difference in means was only marginally significant. For Set 3 (e.g. *baking* vs. *bacon*), the difference of 2 ms (SD = 22) was in the direction predicted by the word-internal PWD hypothesis (mean of stem-suffix = 184 ms, SD = 46; mean of monomorphemic cases = 182 ms, SD = 48) but did not reach significance ( $t = 1.13$ ,  $df = 129$ ,  $p = 0.179$ ). For Set 4 (e.g. *axes* vs. *axis*), where the stem-suffix cases were associated with a 4 ms slower speech rate (see §2.4.5), the mean of the stem-suffix cases was 258 ms (SD = 54) and that of monomorphemic cases 256 ms (SD = 54); the difference of 2 ms (SD = 14) was statistically insignificant ( $t = 0.85$ ,  $df = 31$ ,  $p = 0.202$ ).

The results above, however, do not take into consideration the frequency of the members of each word pair. One-tailed t-test analyses were also conducted for cases (i) in which a suffixed word and its monomorphemic counterpart were equally frequent or (ii) where the frequency of a suffixed word exceeded that of its monomorphemic counterpart. The

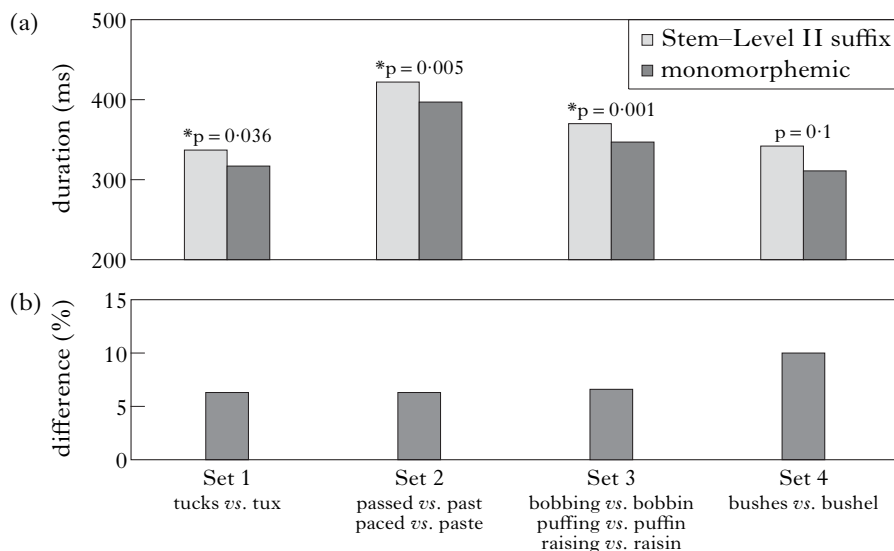
latter pairs provided the most conservative test.<sup>10</sup> The pairs that satisfied the frequency criteria were *tucks vs. tux* in Set 1, *passed vs. past* and *paced vs. paste* in Set 2, *raising vs. raisin*, *puffing vs. puffin* and *bobbing vs. bobbin* in Set 3 and *bushes vs. bushel* in Set 4 (see Fig. 3 for the frequency of each word). For Set 2, the differences between the members of the two pairs turned out to be statistically significant: the rhyme duration of the suffixed words (mean = 263 ms, SD = 31) was on average 2.7% (i.e. 7 ms, SD = 24) longer than that of the monomorphemic words (mean = 256 ms, SD = 32;  $t = 2.10$ ,  $df = 55$ ,  $p = 0.020$ ). Sets 1, 3 and 4, however, did not show any statistically significant differences.

In summary, the normal rate analyses provided little support for the stem-suffix PWD-boundary hypothesis, although the stem-suffix cases in Set 2 were significantly longer than their monomorphemic counterparts in the analysis of pairs satisfying the frequency criteria.

*The slow rate.* At the slow rate, all four sets showed statistically significant differences in the direction predicted by the word-internal PWD hypothesis, i.e. suffixed rhymes were reliably longer than the corresponding stretches in monomorphemic words. For Set 1, the mean of the stem-suffix cases was 336 ms (SD = 54) and that of the monomorphemic cases was 321 ms (SD = 43), a difference of 4.7% (15 ms, SD = 35;  $t = 3.281$ ,  $df = 59$ ,  $p = 0.001$ ). For Set 2, the mean of the stem-suffix cases was 379 ms (SD = 88) and that of the monomorphemic cases was 362 ms (SD = 75), a difference of 4.7% (17 ms, SD = 43;  $t = 2.989$ ,  $df = 59$ ,  $p = 0.002$ ). For Set 3, the mean of the stem-suffix cases was 303 ms (SD = 118) and that of the monomorphemic cases was 289 ms (SD = 109), a difference of 4.8% (14 ms, SD = 46;  $t = 3.432$ ,  $df = 130$ ,  $p = 0.0005$ ). For Set 4, the stem-suffix mean was 409 ms (SD = 145) and the monomorphemic mean was 381 ms (SD = 116), a difference of 7.3% (28 ms, SD = 78;  $t = 2.025$ ,  $df = 31$ ,  $p = 0.026$ ).

Additional analyses were conducted for pairs that satisfied our word-frequency criteria. The mean of the stem-suffix cases was significantly longer than that of the monomorphemic cases in three of the four sets: Set 1 (stem-suffix mean = 337 ms (SD = 43), monomorphemic mean = 317 ms (SD = 53), difference = 20 ms (SD = 46, stem-suffix 6.3% longer than monomorph);  $t = 1.91$ ,  $df = 19$ ,  $p = 0.036$ ), Set 2 (stem-suffix mean = 422 ms (SD = 71), monomorphemic mean = 397 ms (SD = 51), difference = 25 ms (SD = 43, 6.3%);  $t = 3.58$ ,  $df = 39$ ,  $p = 0.0005$ ), Set 3 (stem-suffix mean = 370 ms (SD = 127), monomorphemic mean = 347 ms (SD = 108), difference = 23 ms (SD = 50, 6.6%);  $t = 3.30$ ,  $df = 51$ ,  $p = 0.001$ ). Although Set 4, i.e. the *bushes vs. bushel* pair, also showed the same direction of difference between the two morphological conditions: (stem-suffix mean = 342 ms (SD = 137), monomorphemic mean = 311 ms

<sup>10</sup> The log frequency count of all words used in our experiments is limited to below 2.5, which is estimated not to affect segmental durations (see §2.4.2, with reference to Aylett & Turk's 2004 finding). We nonetheless introduce frequency criteria here, in case a small frequency difference between the members of the same pair should result in an unexpected segmental duration difference.



*Figure 4*

(a) Mean rhyme durations (ms) at the slow speech rate, from pairs satisfying the frequency criteria. P-values preceded by an asterisk indicate statistical significance in a one-tailed t-test comparison of the two morphological conditions ( $\alpha = 0.05$ ). (b) Differences in mean rhyme duration of stem-suffix cases as compared to monomorphemic cases, expressed as a percentage, for pairs satisfying the frequency criteria: (suffix rhyme duration/monomorphemic rhyme duration - 1)  $\times$  100.

(SD = 77)), the difference of 31 ms (SD = 91, 10%) was not statistically significant:  $t = 1.35$ ,  $df = 15$ ,  $p = 0.10$ . This may be due to the fact that there was more inter-speaker variation for the *bushes vs. bushel* pair in Set 4 than for the pairs in other sets, as shown in Fig. 5: only Speaker A showed a substantial difference in the expected direction, 32.4%, while the other three speakers did not.<sup>11</sup> That is, the relatively large difference of 10% in Set 4 comes from Speaker A only, and as a result the difference is non-significant. These results are shown in Fig. 4.

The inter-speaker variation of the other three sets was far smaller, as shown in Fig. 5. For Set 1, four of the five speakers showed the expected difference. For Set 2, the majority of the speakers also showed the expected difference, although Speaker C showed difference in neither of the two item pairs. For Set 3, all speakers except Speaker R showed the expected difference of more than 3% in the comparison of *raising vs. raisin*. For the other pairs (*bobbing vs. bobbin*, *puffing vs. puffin*), there were always two speakers who showed the expected difference of more than 5%. Furthermore, there is one crucial aspect shared by all pairs in the four

<sup>11</sup> Speaker C is not included here. See §2.4.3.

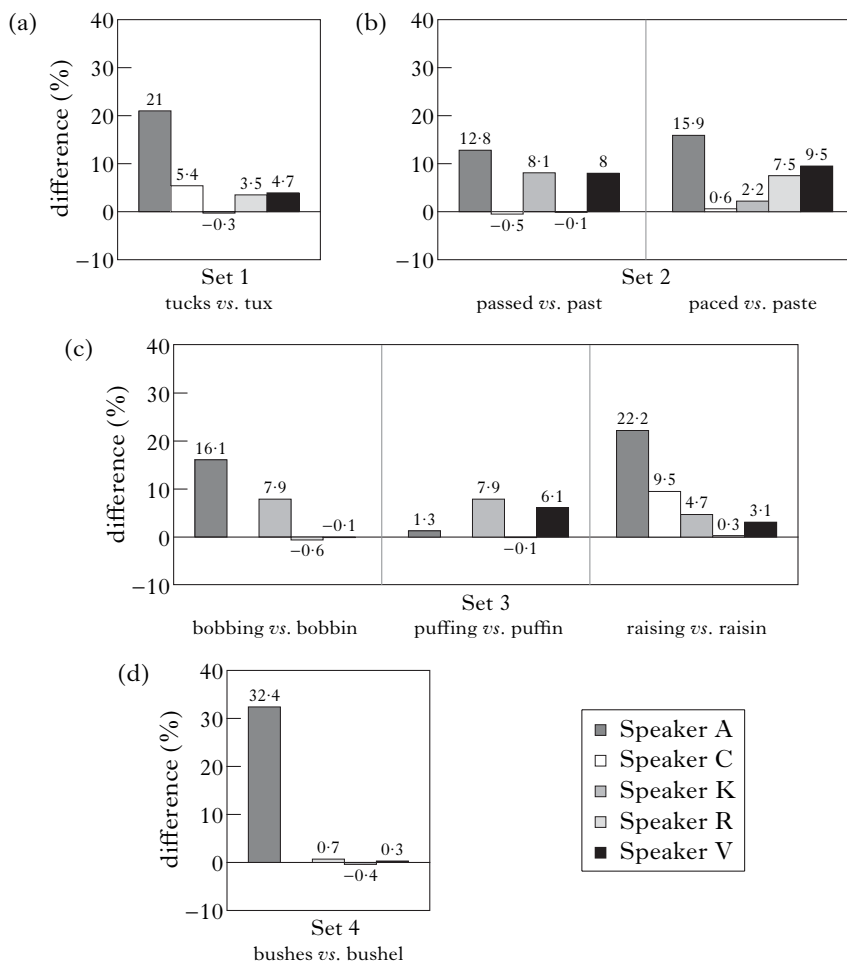


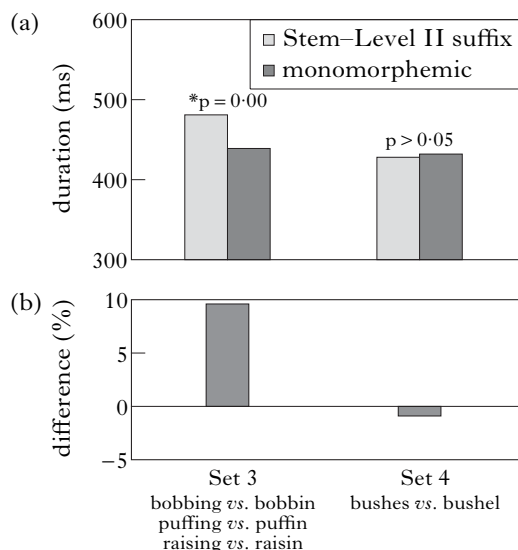
Figure 5

Differences in first syllable mean rhyme durations at the slow rate for stem-suffix *vs.* monomorphemic cases, expressed as a percentage, for pairs satisfying the frequency criteria. Speaker C did not record the *bobbing vs. bobbin*, *puffing vs. puffin* and *bushes vs. bushel* pairs.

sets: when there was a difference of more than 1% between the two morphological conditions, the difference was always in the expected direction.

*The extra-slow rate.* The stimulus pairs in Sets 3 and 4 were also recorded at an extra-slow speech rate. Both sets showed a mean difference in the direction predicted by the word-internal PWd-boundary hypothesis. However, the difference was statistically significant only for Set 3 (stem-suffix mean = 403 ms (SD = 179), monomorphemic mean = 385 ms (SD = 169), difference = 18 ms (SD = 109, 4.7%);  $t = 1.892$ ,  $df =$





*Figure 6*

(a) Mean rhyme durations (ms) at the extra slow speech rate, from pairs satisfying the frequency criteria. P-values preceded by an asterisk indicate statistical significance in a t-test comparison of the two morphological conditions ( $\alpha = 0.05$ ). (b) Differences in mean rhyme duration of stem-suffix cases as compared to monomorphemic cases, expressed as a percentage, for pairs satisfying the frequency criteria.

126,  $p = 0.03$ ). For Set 4, there was no significant difference between the two morphological conditions.

Similar results were obtained for the pairs satisfying our word-frequency criteria in Sets 3 and 4, as shown in Fig. 6. For Set 3, the mean of the stem-suffix cases was 481 ms ( $SD = 195$ ) and that of the monomorphemic cases was 439 ms ( $SD = 182$ ), with a mean difference of 42 ms ( $SD = 67$ , 9.6%;  $t = 4.47$ ,  $df = 50$ ,  $p = 0.00$ ). For Set 4, there was no significant difference between the two morphological conditions.

Although one speaker showed a substantial difference between the two morphological conditions for Set 4 at the slow speech rate, none of the speakers showed a difference at this extra-slow rate. In contrast, four of the five speakers showed a difference of more than 7% in Set 3, as shown in Fig. 7.

**2.4.7 Summary: Experiment 1.** Overall results were consistent with the word-internal PWD hypothesis for most of the word comparisons at the slow speech rate. Analyses for sets of pairs that satisfied the frequency criteria and those that did not satisfy the criteria also showed similar results. Target sequences were longer in suffixed forms for two types of monosyllabic word pairs (e.g. *tacks* vs. *tax* and *missed* vs. *mist*): these

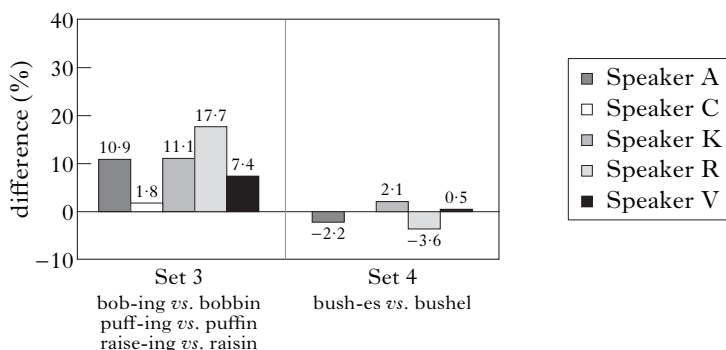


Figure 7

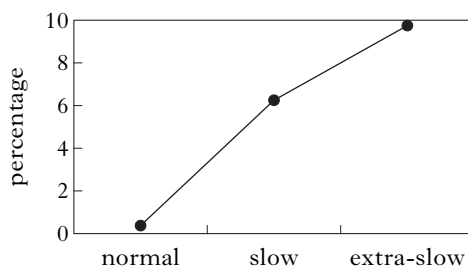
Differences in first syllable mean rhyme durations at the extra slow rate for stem-suffix *vs.* monomorphemic cases, expressed as a percentage, for pairs satisfying the frequency criteria. Speaker C did not record the *bobbing vs. bobbin*, *puffing vs. puffin* and *bushes vs. bushel* pairs.

showed similar behaviour to the *-s* suffixed forms in Schwarzlose & Bradlow (2001). Target sequences were also longer in suffixed forms of disyllabic words controlled for homophony and the number of letters used in the spelling of the target sequences (e.g. *baking vs. bacon*, *puffing vs. puffin*). These results are consistent with the view that the Level II suffixed words are different from their monomorphemic counterparts in their prosodic constituent representation. On this view, the difference in prosodic structure is reflected in surface segment durations through the application of phonetic implementation processes which refer the structure.

It should be noted that there was a difference among sets: although three of the four sets (Sets 1, 2 and 3) showed a difference between Level II suffixed and monomorphemic forms at the slow and extra-slow rates, Set 4 did not. Unfortunately, Set 4 contained just two pairs, and only one of them satisfied our frequency criteria. It is therefore unclear whether the exceptional behaviour of Set 4 is due to the small number of pairs or to the intrinsic nature of *-es* suffixation. This should be further investigated with pairs such as *lex-es vs. lexis* (both of them equally infrequent), *miss-es* (more frequent) *vs. missus* (less frequent), etc.

Given that we did not obtain the expected result in Set 4 for the suffix *-es*, the suffixes that induced the expected difference in our experiment are *-s*, *-ed* and *-ing*. The sets with the two consonantal suffixes, however, contained pairs whose stem-suffix and monomorphemic members were different in the number of letters with which they are spelled, which might have affected the durational differences between these two morphological forms. However, we also obtained longer target durations in stem + *-ing* cases as compared to their monomorphemic counterparts, in spite of the fact that stem + *-ing* cases did not have spelling confounds.

Notice also that the durational differences across the three speech rates in Set 3 are consistent with Beckman & Edwards (1990), who showed that



*Figure 8*

Percentage differences in mean rhyme duration for Set 3 stem-suffix words as compared to monomorphemic words, at different speaking rates.

These differences come from pairs satisfying the frequency criteria.

boundary effects are more exaggerated at slower speech rates. The pairs in Set 3 that satisfy our frequency criteria showed enhancement from the normal to the slow rate, and from the slow to the extra-slow rate. Figure 8 shows how the magnitude of the durational difference between the two morphological conditions of Set 3 varied across the three rates of speech (shown as a percentage), obtained from the pairs that satisfy the frequency criteria (i.e. *bobbing* vs. *bobbin*, *puffing* vs. *puffin*, *raising* vs. *raisin*). There was no such enhancement from the slow rate to the extra-slow rate in Set 4, which was exceptional because only Speaker A showed a substantial difference between the two morphological conditions in the slow rate of speech. The speaker showed an unusually large difference (i.e. 32.4%; see Fig. 5) at the slow rate of speech, while the difference was reduced to -2.2% (see Fig. 7) in the unpredicted direction at the extra-slow rate.

A further finding is that there were inter-speaker differences in the magnitude of the difference between the two morphological conditions. One of our speakers, Speaker A, consistently showed the greatest differences between the two morphological conditions; she was the only speaker that showed the expected difference for Set 4.

## 2.5 Experiment 2: Level I vs. Level II

To confirm the validity of the word-internal PWd-boundary hypothesis, we also compared the durations of stems followed by Level II suffixes with those followed by Level I suffixes. As already discussed in §1.1, the word-internal PWd-boundary hypothesis considered here proposes a PWd boundary at the right edge of stems with Level II suffixes, but not for those with Level I suffixes. This view predicts that stem-final rhyme durations of Level II suffix cases should be longer than those of Level I suffix cases. Experiment 2 tests this prediction.

**2.5.1 Stimulus design.** Stimulus sets consisted of the eight pairs of Level I and Level II suffixed words in Table III.

stem	stem + Level I suffix	stem + Level II suffix
accept	acceptance	accepting
attend	attendance	attending
emerge	emergence	emerging
insist	insistent	insisting
phrase	phrasal	phrasing
assist	assistance	assisting
avoid	avoidance	avoiding

Table III

Item pairs used in Experiment 2.

The Level II suffix tested in this experiment was always *-ing*, while the Level I suffixes were *-ance*, *-ence*, *-ent* and *-al*. The suffixes *-ance* and *-ence* belong to the Level I class because they can trigger stress alternation of the stem (e.g. *ig'nore* vs. *'ignorance*, *pre'fer* vs. *'preference*) and also attach to bound root morphemes, which never stand alone as words (e.g. *brilli-ance* and *experi-ence*). The suffix *-al* is also a Level I morpheme, because it can trigger segmental alternation of the stem (e.g. *n[eɪ]tion* vs. *n[a]tional*) and syllabification of a stem-final sonorant consonant into the following suffix-initial vowel (e.g. *centre* [sɛn.tɹ̩] vs. *central* [sɛn.tɹ̩əl]). All these phenomena support the view that the stems to which these suffixes attach lack independent word status in the lexicon. In this experiment, we chose pairs in which the stem with the Level I suffix (e.g. *attend* in *attendance*) and that with the Level II suffix (e.g. *attend* in *attending*) are homophonous.

The frequency data for these words are shown in Fig. 9. For all pairs except for *assisting* vs. *assistance*, the log frequency of the Level II suffixation cases was either greater than or the same as that of the Level I suffixation cases. These cases therefore satisfy our frequency criteria.

**2.5.2 Speakers.** Three of the native female speakers of Scottish Standard English who participated in Experiment 1 served as paid subjects.

**2.5.3 Elicitation and recording.** Only slow speech was elicited in Experiment 2, given that the results from Experiment 1 suggested that the durational effects of the stem–Level II suffix boundaries were greatest and most reliable at the slow rate. Four repetitions of each pair were recorded, giving a total of 168 items (i.e. 3 speakers × 4 repetitions × 2 morphological conditions × 7 pairs). The rest of the recording procedure was the same as that of Experiment 1. The carrier sentences in which each target word was embedded are shown in the Appendix.

**2.5.4 Measurements and analyses.** We measured the rhyme duration of the stem-final syllable of each target word, e.g. /ɛnd/ in *attendance* and

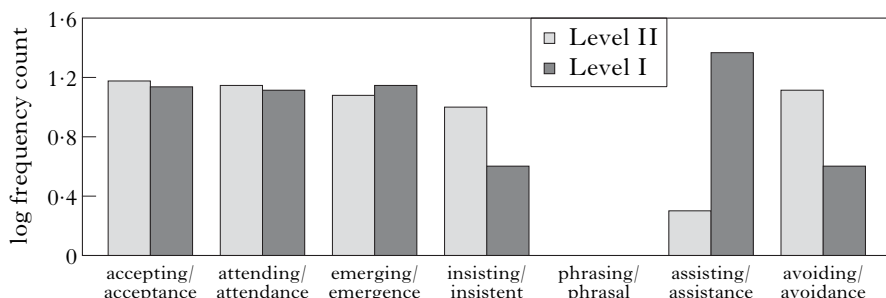


Figure 9

Log word frequency count of items used in Experiment 2.

*attending*. Measurement criteria were the same as those for the previous experiment: the target sequence was measured as the interval between the onset of constriction release of the first consonant and that of the coda consonant of the stem-final syllable (see again note 11). For the target sequence with a sonorant onset [ɹ], i.e. [ɹeɪz] in *phrasal* and *phrasing*, it was impossible to determine a boundary between the sonorant consonant and the following vowel. We therefore included the sonorant onset consonant in the target measurement sequence. Part of the initial word of each carrier sentence (*Say* or *Type*) was measured as a check for overall speech-rate differences (see §2.4.5 for more detailed discussion). Due to segmentation difficulties, 17 item pairs were discarded, giving 67 pairs out of 84 for analysis. 14 pairs discarded out of 17 were Speaker K's, due to the unclear onset of the stem-final consonant release. Data were analysed using one-tailed paired t-tests; each repetition of a test sentence in one morphological condition was always paired with that of the other morphological condition within a single block.

**2.5.5 Results.** A one-tailed paired t-test was carried out on the control-segment durations, i.e. the durations of the rhymes of *Say* and *Type*, which immediately preceded the target words and is an indicator for overall speech rate. No statistically significant differences in the control-segment durations between the Level II cases and the Level I cases were observed; we therefore found no evidence of any speech-rate difference between the two morphological conditions.

The mean rhyme duration of Level II suffix cases was 7.2% (i.e. 30 ms, SD = 76) longer than that of the Level I suffix cases, this difference was statistically significant, according to a one-tailed paired t-test: the mean of the Level II suffix cases was 445 ms (SD = 112) and that of the Level I suffix cases was 415 ms (SD = 105;  $t = 3.205$ ,  $df = 66$ ,  $p = 0.001$ ). We also ran an analysis without the *assisting* vs. *assistance* pair that violated our frequency criteria, only to obtain the same result: the Level II suffix cases (mean = 446 ms, SD = 113) were significantly longer (8%, 33 ms, SD =

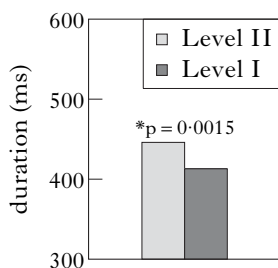


Figure 10

Mean first syllable rhyme durations (ms) at the slow speech rate for Level I *vs.*

Level II suffixed stems, from pairs satisfying the frequency criteria. The asterisk indicates statistical significance in paired t-test comparisons ( $\alpha = 0.05$ ).

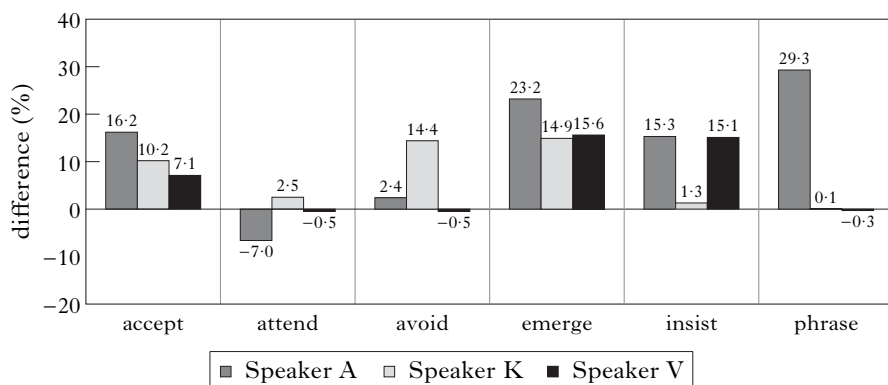


Figure 11

Percentage differences in mean first syllable rhyme duration for Level II suffixed *vs.* Level I suffixed forms at the slow rate. All pairs whose means are shown here satisfy the frequency criteria.

81) than the Level I suffixes (mean = 413 ms, SD = 102;  $t = 3.055$ ,  $df = 57$ ,  $p = 0.0015$ ). This is shown in Fig. 10.

*Speaker and pair differences.* Figure 11 shows pair-by-speaker mean values of the differences. The pairs presented all satisfy the frequency criteria. Three of the six pairs (i.e. *accepting vs. acceptance*, *emerging vs. emergence*, *insisting vs. insistence*) showed the expected difference for all speakers. There was always at least one speaker who showed the expected difference for the other three pairs. That is, there was no instance in which all three speakers unanimously showed a difference in the unexpected direction. Furthermore, when there was a difference of more than 1%, it was always in the predicted direction, except for the *attend-* pair as produced by Speaker A.

Thus, although there was some pair-by-pair and speaker-by-speaker variation, the overall tendency was that the rhyme duration in Level II suffix cases was longer than that in Level I suffix cases.

**2.5.6 Summary: experiment 2.** Our results are consistent with the view that the stem–Level II suffix words differ from their stem–Level I suffix counterparts in their prosodic constituent representation. That is, Level II suffixes adjoin to a word-level constituent, and the presence of the word boundary can result in a longer stem duration at a slow rate of speech. In contrast, the shorter stem-final rhyme durations before Level I suffixes suggest that no such word boundary is present between the stem and the following Level I suffix.

## **2.6 Another possible account: phonetic paradigm uniformity**

The results obtained in Experiment 1 and Experiment 2 are in conflict with Raffelsiefen's (2005) position that word-boundary phenomena observed before vowel-initial Level II suffixes should be accounted for by phonological paradigm uniformity rather than the presence of a PWd boundary. The results are also inconsistent with a serialist approach in which word-internal boundaries are erased before the final surface form is derived, due to bracketing erasure (Kiparsky 1982).

There is, however, another possible account for the results: the *phonetic* paradigm uniformity approach originally suggested by Steriade (2000).<sup>12</sup> According to phonetic paradigm uniformity, not only phonological features but also non-contrasting phonetic details, such as consonant closure duration, are transferred from a base form to related forms. For example, in Steriade's study, the duration and other phonetic details of French [d] and [ʁ] in a fully articulated form without any deletion of schwa [də.ʁol] *de rôle* and in a form in which the schwa undergoes deletion [dʁol] *d'rôle* are identical, in spite of the fact that they do not share the same syllable structure. The fully articulated form in this case is the base, and its phonetic details are transferred to the related form with schwa deletion.

The same analysis would in principle be available for the cases discussed here: durational characteristics of the independent word *bake* as a base would be extended to the related word *baking*. The rhyme [ek] in the monosyllabic independent word *bake* is often longer than in disyllabic *bacon*, because of word- (or phrase-) final lengthening and/or lack of polysyllabic shortening. In Turk & Shattuck-Hufnagel (2000), for example, the difference between them was about 30 ms (10%) in phrase-medial position with a focal pitch accent at normal speech rate, and even larger in phrase-final position. In the phonetic paradigm uniformity theory, the phonetic properties of *bake*, including the duration of [ek], would be transferred to its related form *baking*.

<sup>12</sup> We are grateful to José Hualde for suggesting this possibility.

The same mechanism would account for the Level I *vs.* Level II comparisons as well: words consisting of a stem and a Level II suffix such as *attending* could be associated with the base form *attend* in the mental representation, while words consisting of a stem and a Level I suffix such as *attendance* could be represented independently of the base form. The durational properties of *attend* would therefore be transferred to its related form *attending*, but not to the independent form *attendance*.

This phonetic paradigm uniformity approach faces an obvious problem, however. The duration of the base form (e.g. of the monosyllabic word *bake*) varies across different prosodic positions in phrases and sentences: its word-final rhyme is longer at the right edge of a prosodic unit than in a medial position of the same unit (e.g. it is longer in the final position of an intonational phrase than in intonational phrase-medial position, and it is longer in the final position of a phonological phrase than in phonological phrase-medial position, and so on; Wightman *et al.* 1992). In addition, segments in monosyllabic words vary with respect to their phrasal prominence: phrasally stressed (i.e. pitch accent-bearing) syllables are known to be longer than non-phrasally stressed syllables. Such durational variation of words and syllables in different prosodic positions is well documented not only for English (Cooper & Paccia-Cooper 1980, Price *et al.* 1991, Wightman *et al.* 1992), but also for many other languages (French: Fougeron & Keating 1997; Dutch: Cambier-Langeveld 2000; Korean: Cho & Keating 2001). Finally, segment, syllable and word durations vary with respect to rate of speech.

All of these effects make it extremely complicated to specify the duration of a supposed base form, and hence to formalise the output–output correspondence between the durational properties of the base form and those of related forms. In order for the durational identity of two related representations to be achieved by paradigm uniformity, low-level durational phonetic details which are infinitely variable across different prosodic positions, prominence locations and speech rate must be represented in surface phonological output representations. Although it would be possible to formulate a phonetic paradigm uniformity theory to account for our results, it would significantly complicate the theory of phonological representation.

Empirically speaking, too, the duration of the suffixed stem observed in our study was much smaller than that of possible base forms: even a supposed base form that occurs in a phrase-medial position is likely to be longer than the suffixed stem observed in our study. In our current experiments, we found that the mean rhyme duration of suffixed stems (e.g. [ek] in *baking*) was only 2 ms longer than that of the corresponding monomorphemic words (e.g. [ek] in *bacon*) at a normal speech rate. In contrast, results from an earlier experiment, reported in Sugahara & Turk (2004), show that the rhyme duration of unsuffixed words (i.e. the supposed base forms for the suffixed stem) in phrase-medial position (e.g. [ek] in *Bake Enforce*) was 10 ms longer on average than that of monomorphemic words (e.g. [ek] in *Bacon Force*) at a normal speech rate. This



indicates that the greater length of a suffixed stem as compared to its monomorphemic counterpart cannot be straightforwardly attributed to paradigm levelling of duration patterns between the base forms and the suffixed stems.

In contrast, the word-internal PWd-boundary hypothesis predicts that the difference in duration between the suffixed stem and its monomorphemic counterpart should be far subtler than that of the corresponding phrase-medial independent word. This prediction of the PWd-boundary hypothesis comes from two aspects of the prosodic representation of suffixed words: (i) relative boundary strength and hierarchical level of the word-internal PWd and (ii) the number of PWd boundaries intervening between the stem and the Level II suffix.

With respect to the relative strength of the stem boundary, the prosodic word directly dominating the stem,  $PWd_{Stem}$ , is always embedded in a higher prosodic word dominating the whole suffixed word,  $PWd_{Wd}$ . In contrast, a phrase-medial independent word forms an independent  $PWd_{Wd}$ . Therefore, by analogy with other phonetic implementation processes such as domain-final lengthening that correlate with the level of prosodic boundaries in the hierarchy, we would expect smaller effects at the relatively lower  $PWd_{stem}$  boundary than at the phrase-medial  $PWd_{Wd}$  boundary.

With respect to the number of prosodic word boundaries between the stem and the following suffix,  $PWd_{Stem}$  is followed by a suffix that is immediately dominated by  $PWd_{Wd}$  without constituting its own PWd. Therefore only a single PWd edge, the right edge of  $PWd_{Stem}$ , intervenes between the stem and the following suffix (see Fig. 2). In contrast, two PWd boundaries (i.e. both the right edge of the preceding word and the left edge of the following word) always intervene between two words in phrase-medial position (see again Fig. 2). Assuming that the magnitude of polysyllabic shortening attenuation that operates on a sequence of multiple syllables is positively correlated with the number of PWd boundaries between syllables (see §2.2 for more discussion on this issue), we expect less shortening of syllables in an independent word in phrase-medial position than of syllables in a stem followed by a suffix. In this way, the PWd-boundary hypothesis provides a straightforward account of the subtler durational effects of proposed  $PWd_{Stem}$  as compared to  $PWd_{Wd}$ .

It is important to note that there is another possible hypothesis consistent with our data that involves a word boundary immediately before the Level II suffix. It is possible to formalise the effect we found in terms of the presence of a morphological word boundary immediately before the Level II suffix, where the phonetic implementation rules directly refer to the morphological word boundary. This alternative direct reference hypothesis has been adopted in some previous studies (e.g. Cho 2001), and our experimental results do not at this point provide any empirical evidence to favour our prosodic structure (i.e. indirect reference) hypothesis over the direct reference hypothesis. Nonetheless, as already mentioned in §2.1.2, there is a considerable body of evidence at the phrasal level for the

non-isomorphism between prosodic constituent structure that has led many people to postulate that prosodic, rather than morphosyntactic, structure is the direct determinant of phonetic behaviour (see e.g. Ferreira 1993).

In summary, the prosodic word-boundary hypothesis provides a more straightforward account for the data than the phonetic paradigm uniformity theory. To be equally viable, the latter would at least need to be modified in order to give a principled way of choosing appropriate base forms, and to account for differences in duration between base and suffixed forms.

### 3 Is there durational evidence for the stress-delimited word-internal foot?

It has been widely assumed that word-internal feet constitute a level of constituent structure intermediate between the level of the syllable and the PWd (Hayes 1980, Selkirk 1980a, Nespor & Vogel 1986, among others). Developing the notion of binary branching metrical constituent structure, and eliminating the feature  $[\pm\text{stress}]$  proposed by Liberman & Prince (1977), Hayes (1980) and Selkirk (1980a) introduce the stress-delimited trochaic foot as a word-internal constituent responsible for lexical stress assignment in English. In this approach, the word-internal foot is the domain in which exactly one lexical stress is assigned; the stressed syllable may be followed by an unstressed syllable or syllables in the same foot. Each word-internal foot is assumed to be at least binary in terms of syllables or moras (Mester 1994, Hayes 1995).

Word-internal feet have been also used to account for interactions between stress placement and syllable weight (Selkirk 1980a), the truncation of unstressed syllables by infants (e.g. Gerken 1994a, b) and infixation (McCarthy 1982, Hammond 1999). However, as discussed below, most of the facts here can be explained without postulating word-internal feet. In our experiment, we compared the durations of syllables in monosyllabic feet, e.g. [am] in ('am)(bush), with comparable syllable durations in disyllabic feet, e.g. [am] in ('am.ber), on the assumption that syllables belonging to monosyllabic feet in two-stress words should be longer than syllables belonging to disyllabic feet in single-stress words, due to polysyllabic shortening or pre-boundary lengthening, as discussed above.<sup>13</sup>

<sup>13</sup> The presence or absence of lexical stress in English is diagnosed by the surface segmental and suprasegmental characteristics of each syllable. Surface segmental stress diagnostics include the vowel quality of syllable nuclei and the allophony of onset consonants. Stressed syllables are realised with full vowels (except for dialects with stressed syllabic liquid consonants as in *fur* and *pull*), and their onset voiceless stop consonants are aspirated.

### 3.1 Arguments for word-internal feet in English

In this section, we discuss phonological and phonetic phenomena often cited as supporting evidence for word-internal feet. They can, however, be accounted for in alternative approaches, based on notions such as the bimoraicity of stressed syllables, word-final lengthening and native speakers' preference for strictly alternating rhythmic patterns between strong and weak syllables.

**3.1.1 *Stress placement and syllable weight in English.*** One of the motivations for the word-internal foot constituent in English comes from the interaction between the location of lexical stress and syllable weight (Selkirk 1980a). In American English, open syllables with a stressed lax (i.e. short monomoraic) vowel are rare or non-existent (a) in word-final position: \* $[\dots C'\check{V}\#]$ , and (b) in word-medial position immediately followed by another stressed syllable when at most one consonant intervenes between the lax vowel and the following stressed vowel: \* $[\dots C'\check{V}(C)'VC \dots]$ .<sup>14</sup>

This fact can be accounted for by foot theory in the following way. Since every stressed vowel, whether short (monomoraic) or long (bimoraic), must be the head of a left-headed foot, all word-final and word-medial stressed syllables must be foot-initial. Because they are not followed by any unstressed syllables, they must form a foot on their own. A problem arises when these syllables are monomoraic, because a foot consisting of only a monomoraic syllable is illegal if each foot should dominate at least two syllables or moras (Mester 1994). American English therefore has neither a word-final nor a word-medial lax stressed vowel immediately followed by another stressed syllable (with at most one intervening consonant). In this way, the foot hypothesis gives a simple account of the interaction between the location of a stressed syllable and its weight.

Foot theory, however, is not the only possible account for these facts. First, the lack of word-final lax vowels is accounted for by the familiar requirement that a stressed syllable should be bimoraic. Secondly, the lack of a lax vowel immediately followed by only one consonant and a stressed vowel is also accounted for by the bimoraicity requirement of stress, and a requirement for an onset consonant in every stressed syllable.<sup>15</sup> The \* $[\dots C'\check{V}C'VC \dots]$  configuration violates one of the requirements no matter what the syllable affiliation of the intervening consonant is. If it is affiliated to the initial syllable, then the following syllable violates

<sup>14</sup> Exceptionally, word-initial stressed syllables containing lax vowels may be open, as in *satire* [ $sæ,^{th}aɪ$ ] and *rabbi* [ $ˈɹæ,baɪ$ ]. Accounting for these forms requires a principle that allows open lax syllables in a word-initial syllable bearing word-level prominence, regardless of syllable weight.

<sup>15</sup> These requirements are captured as positional markedness by Smith (2002) in the framework of Optimality Theory.

the onset requirement. If it is affiliated to the second syllable, then the preceding syllable with a lax vowel violates the bimoraicity. As a result, the configuration fails to surface.

Thus the lack of syllables ending with a lax vowel in word-final and pre-tonic positions can be explained by syllable-related constraints or principles which do not refer to foot structure.

**3.1.2 Unstressed syllable truncation.** Some additional pieces of evidence for the foot hypothesis in English come from unstressed syllable omissions in infant speech: speakers often truncate unstressed syllables that are not parsed into a trochaic stress (SW) foot. Gerken (1994a, b), Wijnen *et al.* (1994) and Demuth (1996), among others, claim that infants make use of a trochaic rhythmic unit, i.e. the stress foot, as a template for word production. For example, children often produce 'raffe for *gi'raffe*, while keeping both syllables in a strong-weak sequence like 'tiger, and produce 'nana for *ba'nana*, with omission of the first syllable. Gerken (1994b) further investigates infants' imitation tasks for four-syllable words, and found that children often truncated even non-initial unstressed syllables when these syllables did not fit in the foot template, i.e. unstressed syllables not immediately following a stressed syllable. Since the foot template allows at most two syllables (i.e. a stressed syllable and an immediately following unstressed one), such an unstressed syllable is unparsed and undergoes truncation.

The foot-based account, however, is not the only way to capture the truncation phenomena. First, the tendency for word-initial unstressed syllable truncation may be due to the fact that the majority of English words start with a stressed syllable (Cutler & Butterfield 1992). That is, infants' unstressed syllable truncation reflects the distribution of stress patterns in the words they hear, and they adjust their output to be similar to more familiar initially stressed words by omitting the word-initial unstressed syllable. In addition, word-initial unstressed syllables may be less perceptually salient than word-final unstressed syllables in disyllabic words in some contexts, due to the higher probability of phrase-final lengthening, and accentual lengthening of the preceding accented syllable carries over to word-final syllables (Wightman *et al.* 1992, Turk & White 1999). Word-medial unstressed vowel truncation may reflect a preference for stress patterns that have a strict alternation between strong and weak syllables (i.e. strong-weak-strong-weak) to more anomalous patterns such as two consecutive weak syllables followed by a strong syllable, etc. In addition, children's tendency to truncate a word-medial unstressed syllable not immediately preceded by a stressed syllable may also be due to such a syllable not being as perceptually salient as an unstressed syllable immediately preceded by a stressed syllable (in phrasally stressed words), because of the rightward spread of phrasal stress-related (accentual) lengthening (Turk & Sawusch 1997, Turk & White 1999). It is therefore possible to explain most unstressed syllable omissions in English without recourse to the word-internal foot.

3.1.3 *Expletive infixation*. Another argument provided by previous studies for the sublexical foot constituent hypothesis in English comes from ‘expletive infixation’, discussed in detail by McCarthy (1982) and Hammond (1999). Although this phenomenon has also been treated as evidence for the English sublexical foot, other possible accounts are also available: these invoke native speakers’ preference for alternating strong–weak rhythmic patterns.

English infixation is a productive morphological process in which expletives such as *fuckin’*, *goddam* (American English) and *bloody* (British English) are inserted in a word-medial position to give an emphatic connotation to words, e.g. *fan-fuckin-tastic* (derived from *fantastic*). McCarthy and Hammond claim that infixes are never able to break up a foot. Therefore, one never finds infixed words like *\*a('mal-<sub>1</sub>fuckin-ga)* (*mated*), where the infixation breaks up two syllables that are parsed into the same trochaic foot, whereas a stressed syllable and an immediately preceding unstressed syllable are easily separated, as in *a('malga-<sub>1</sub>fuckin-)* (*mated*), (*Kala*)*ma-<sub>1</sub>fuckin-('zoo)*, (*Kala*)-*fuckin-ma('zoo)* and *ad-<sub>1</sub>fuckin-('vance)*, because the infixation does not break up two syllables within the same foot.

Although this foot-based formalisation of English infixation is simple and straightforward, it is also possible to formalise the phenomenon by stating that native speakers of English prefer the rhythmic pattern to alternate strictly between strong and weak syllables. The strong–weak alternation pattern by itself is a simple and primitive notion, which does not necessarily require a foot-level constituent in its definition. The English infixes all consist of an initial strong syllable followed by a weak syllable as in *fuc* (S) *kin* (W), and infixation immediately after a stressed (strong) syllable is therefore undesirable because such infixation would yield two consecutive strong syllables, thus violating the favoured alternation pattern. Therefore *\*a'mal-<sub>1</sub>fuckin-ga<sub>1</sub>mated* (WS-SW-WSW) is bad because the infixation creates a stress clash between *'mal* (S) and *₁fuc* (S). In contrast, *a'malga-<sub>1</sub>fuckin-<sub>1</sub>mated* (WSW-SW-SW) is preferred, because it avoids both stress clash and two consecutive unstressed syllables. However, stress-clash cases like *₁fan-<sub>1</sub>fuckin-<sub>1</sub>tastic* (S-SW-SW) are allowed because no better option is available: if the infix is inserted between the penultimate stressed syllable and the final unstressed syllable as in *\*₁fan'tas-<sub>1</sub>fuckin-tic* (SS-SW-W), the outcome representation contains two stress clashes (i.e. *₁fan'tas* and *'tas-<sub>1</sub>fuc...*), while the well-formed representation *₁fan-<sub>1</sub>fuckin-<sub>1</sub>tastic* (S-SW-SW) has only one. Thus the foot-based account is not the only possible explanation for infixation in English.

3.1.4 *Durational evidence for word-internal feet?* Phonetic evidence for word-internal feet in English is ambiguous at best. Although some results reported in the literature, e.g. Fowler (1981) and Rakerd *et al.* (1987), are consistent with the view that word-internal feet affect segmental duration, White (2002) suggests that these results are more likely to be word-sized

constituent-related phenomena, rather than phenomena relating to word-internal feet. Van Lancker *et al.* (1988) also presents effects less easily interpreted without recourse to the foot, but the observed effects have a variety of experimental flaws which need to be corrected in subsequent studies.

Fowler (1981) found asymmetric polysyllabic shortening between trochaic (strong–weak) words and iambic (weak–strong) words in a metronome-paced speech-production experiment of nonsense words in English. Stressed syllables immediately followed by unstressed syllables in trochaic words such as /'sisa/ were shorter on the average by 101 ms than stressed syllables produced in isolation in the word /'si/. However, stressed syllables immediately preceded by unstressed syllables in iambic words such as /s<sub>A</sub>'si/ were only 9 ms shorter than those produced in monosyllabic words. Rakerd *et al.* (1987) interpret this asymmetric polysyllabic shortening as evidence for the trochaic foot hypothesis. They consider the domain of polysyllabic shortening to be the trochaic foot constituent, while the stressed syllable of iambic words forms a monosyllabic foot of its own, and therefore does not undergo within-foot polysyllabic shortening. However, White's (2002) research on the durational patterns of phrases containing identical foot structure, e.g. [(*'shake*)<sub>Ft</sub> (*down*)<sub>Ft</sub>]<sub>Wd</sub> [(*'stairs*)<sub>Ft</sub>]<sub>Wd</sub> 'shake-down stairs' *vs.* [(*'shake*)<sub>Ft</sub>]<sub>Wd</sub> [(*down*)<sub>Ft</sub> (*'stairs*)<sub>Ft</sub>]<sub>Wd</sub> 'shake down-stairs' – each foot in those words is non-branching, consisting of a stressed syllable only – suggests that this asymmetric behaviour is not due to foot structure. The duration of *shake* was shorter in *shakedown stairs* than in *shake downstairs*, indicating that the asymmetry had to do with word structure rather than foot structure, while word-final syllables were consistently longer than non-final syllables. White's results suggest that asymmetric shortening of syllables has nothing to do with the number of syllables or the position of syllables in a foot, but instead relates to processes involving word-sized constituents, such as word-final lengthening or polysyllabic shortening (see Turk & Shattuck-Hufnagel 2000 for a discussion of these mechanisms).

Van Lancker *et al.* (1988) also provide durational observations that support the foot hypothesis. However, we show below that their study has several flaws, and as a result their observations do not offer unambiguous support for the word-internal foot hypothesis. They found that syllable duration in English varied with the quality of the immediately following syllable: it was about 20 ms longer when immediately followed by a stressed syllable, e.g. [an] in '*An**des*', than when followed by an unstressed syllable, e.g. [an] in '*Andy*'s. These results can be accounted for by foot-level polysyllabic shortening, if one assumes that a foot boundary separates the two heavy syllables in the strong–strong word, (*An*)(*des*), but that the two syllables in the strong–weak word belong to the same foot, (*Andy*'s). However, several factors confounded their results. For example, their comparison pairs differed in syllable structure (e.g. *to.ma.to vs. tor.ture*), syllable count within each test word (e.g. '*complex vs. complicated, vortex vs. vortices*'), the level of stress prominence (e.g. *ban'dana*

(secondary stress) *vs.* 'bandit (primary stress)), position within carrier sentences (e.g. *Nimrod* in utterance-initial position *vs.* *nimble* in utterance-medial position) and the mixing of proper names and common nouns (e.g. *Sunday vs. sundown*), etc. Further controlled experiments are therefore required to provide durational evidence for the word-internal foot hypothesis.

In sum, no unambiguous phonetic or phonological evidence has yet been provided in support of the word-internal foot. However, the potential for durational support exists, if Van Lancker *et al.*'s observation of foot-level polysyllabic shortening holds even when confounding factors are removed. In the experiment presented below, we control as many of these factors as possible, in order to see if foot-level polysyllabic shortening persists, as predicted by the word-internal foot hypothesis.

### 3.2 Experiment 3

3.2.1 *Materials.* Materials consisted of seven pairs of (a) disyllabic words consisting of a primary stressed syllable followed by a secondary stressed syllable, e.g. 'Ar.gos, where the first syllable constitutes a monosyllabic foot, i.e. ('Ar)<sub>Ft</sub> (gos)<sub>Ft</sub>, and (b) disyllabic words consisting of a primary stressed syllable followed by an unstressed syllable, e.g. 'Argus, where the first syllable belongs to a disyllabic foot, i.e. ('Ar.gus)<sub>Ft</sub> (see Table IV).

two monosyllabic feet	one disyllabic foot
('Con)(cord)	('conquer)
('com)(mune)	('common)
('cli)(max)	('climate)
('cent)(taur)	('centre)
('a)(corn)	('acre)
('am)(bush)	('amber)
('Ar)(gos)	('Argus)

*Table IV*

Item pairs used in Experiment 3.

The segmental organisation of the initial target syllable, i.e. the primary stressed syllable, was kept constant across stress conditions. The vowel quality of the second syllable was different in the two stress conditions, because stressed syllables in English are by definition realised with full vowel nuclei, while unstressed vowels are not (see note 13). In our materials, as shown in the list of item pairs in Table IV, unstressed second syllables always had either schwa, as in *climate*, or a syllabic consonant,

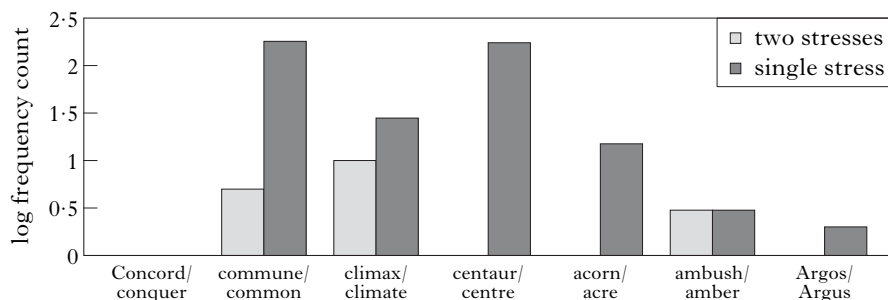


Figure 12

Log word frequency count of items used in Experiment 3.

as in *amber* or *common*, as their nucleus. The log word frequencies of the stimulus pairs are given in Fig. 12.

These items were embedded in carrier sentences and preceded by background sentences designed to elicit primary phrasal stress on the target word, as in (12) (see the Appendix for the full list).<sup>16</sup>

(12) To overcome and take control. Say '*conquer*' for me.

**3.2.2 Speakers and recordings.** The same five speakers who participated in Experiment 1 read the stimulus set in this experiment. Recording procedures were the same as those in Experiment 1.

**3.2.3 Measurements.** The target sequence in this experiment consisted of the rhyme of the first syllable in disyllabic words containing two stresses, as well as the corresponding stretch in the words containing single stress. The onset of the target stretch was the onset-consonant release, when an onset was present, thereby including any onset VOT in the duration of the target sequence. For the pair *climax* vs. *climate*, where there is an obstruent-sonorant consonant cluster [kl] in onset position, we took the obstruent-consonant release to be the onset of the rhyme: the rhyme therefore included VOT as well as the sonorant onset consonant [l]. The offset of the rhyme differed for different stimulus pairs. When only

<sup>16</sup> One of the reviewers asked whether on *for* in the carrier (*Say \_\_ for me*) bore stress: if *for* was unstressed, there would be a possibility that it would be incorporated into a prosodic word and a foot of the preceding target word. We checked the vowel quality of *for* produced by every speaker: four of the five speakers constantly produced the nucleus vowel of *for* as a full vowel [ɔ] (i.e. a stressed vowel). For these four speakers, we could guarantee the presence of a foot boundary between the target word and the following *for*. The other speaker (Speaker R) produced the vowel of *for* as a reduced (unstressed) vowel. She nonetheless consistently inserted an audible pause between the preceding target word and the following *for*, so we could safely conclude that a prosodic word and a foot boundary were present between those two words.



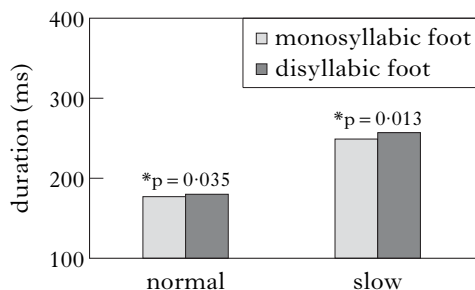


Figure 13

Mean first syllable rhyme durations (ms) at the slow rate for monosyllabic and disyllabic foot cases. The asterisks indicate statistical significance in paired t-test comparisons ( $\alpha = 0.05$ ).

a single consonant was present between the two syllables (e.g. *acorn vs. acre*), the offset of the initial syllable nucleus F2 energy was taken as the offset of the rhyme. When there was a consonant cluster between the first and second syllable nuclei ([ŋk] in *Concord vs. conquer*, [mb] in *ambush vs. amber*, [nt] in *centaur vs. centre*, [ɹg] in *Argos vs. Argus*), the rhyme included acoustic evidence of the first member of the consonant cluster. For homorganic clusters, the offset of the rhyme was taken to be the offset of nasal formant energy. Cases of unclear boundaries between the two consonants were excluded from our analyses (normal rate: four pairs of *Concord vs. conquer*, ten pairs of *ambush vs. amber*, eight pairs of *centaur vs. centre*; slow rate: five pairs of *ambush vs. amber*, four pairs of *centaur vs. centre*).

**3.2.4 Results.** One-tailed paired t-tests were performed to assess whether the mean rhyme duration of the monosyllabic foot was significantly longer than that of the disyllabic foot. As shown in Fig. 13, we found that, contrary to Van Lancker *et al.*'s (1988) results, the mean rhyme duration of the monosyllabic feet, e.g. [ɑɪ] in ('Ar)(gos), was slightly shorter than that of the initial syllable rhyme duration of the disyllabic feet, e.g. [ɑɪ] in ('Argus). At the normal rate of speech, the mean duration of the monosyllabic foot rhymes was 177 ms (SD = 44), while that of the rhymes of the disyllabic feet was 180 ms (SD = 43), which means that the rhyme of the monosyllabic foot was about 2% (3 ms, SD = 18) shorter than that of the disyllabic foot, contrary to our prediction. This difference was statistically significant according to a one-tailed paired t-test ( $t = -1.821$ ,  $df = 110$ ,  $p = 0.035$ ). At the slow rate, the mean duration of monosyllabic foot rhymes was 249 ms (SD = 72), while that of the stressed syllable rhyme in a disyllabic foot was 257 ms (SD = 75), i.e. the former was about 3% (8 ms, SD = 38) shorter than the latter. The difference was again statistically significant according to a one-tailed paired t-test ( $t = -2.274$ ,  $df = 127$ ,  $p = 0.013$ ).

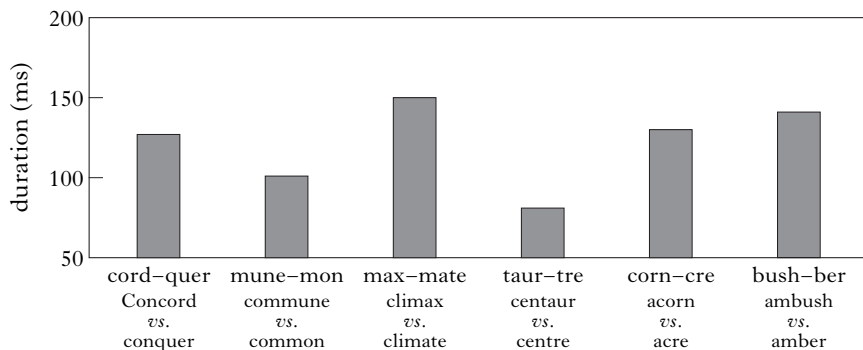


Figure 14

Mean difference (ms) in second syllable duration for monosyllabic and disyllabic foot cases at the slow rate, obtained by subtracting the latter from the former.

**3.2.5 Partial effects.** A possible criticism of this set of comparisons would be that some pairs differed in their surface segment count. For example, the second syllable of the monosyllabic foot case (*'cli*)(*max*) consists of four segments [maks], while that of the disyllabic foot case (*'climate*) consists of three segments [mæt]. Six of our seven pairs actually came with such second syllable segment-count differences: the monosyllabic feet cases contained more second syllable segments than the disyllabic foot cases. Crude measurements of the second syllables of those six word pairs (including final coda consonant release intervals) at the slow rate of speech also show that the second syllable duration of the monosyllabic foot cases was 80 ms to 150 ms longer than that of the disyllabic foot cases, as shown in Fig. 14.

Given this, one might speculate that the shorter rhyme duration of the initial syllable in the monosyllabic feet cases observed above would be due to word-level polysegmental shortening, i.e. compensatory shortening of the initial syllable under pressure from the heavier second syllable, and that this effect could have obscured the foot-boundary effect.

It is therefore necessary to analyse separately the only pair without such second syllable segment-count differences, i.e. (*'Ar*)(*gos*) vs. (*'Argus*).<sup>17</sup> The second syllable rhyme-duration differences between the two foot conditions in this pair turned out to be relatively small in comparison to

<sup>17</sup> The CELEX database does not contain *Argos* in its word list, and therefore we could not obtain its frequency. *Argos*, however, is a familiar term among people in the United Kingdom, because it is the name of a popular catalogue/online shopping store. Therefore, we consider the actual frequency of the word *Argos* to exceed that of its counterpart *Argus*, because *Argus* is a figure in Greek mythology that is not necessarily used on a daily basis.

those in the six pairs above: only 13 ms to 22 ms (at the slow rate of speech).

However, the results of the separate analysis of the rhyme durations of the first syllables of this pair did not support the foot hypothesis either. The mean duration of the initial rhyme of ('*Ar*)(*gos*) was 164 ms (SD = 19) at the normal rate of speech and 249 ms (SD = 41) at the slow rate of speech, while that of ('*Argus*) was 165 ms (SD = 17.8) at the normal rate of speech and 255 ms (SD = 39) at the slow rate of speech. That is, the mean of the initial rhyme duration of ('*Ar*)(*gos*) was again shorter than that of ('*Argus*), contrary to our prediction. The difference was, however, statistically insignificant under a one-tailed paired t-test (normal rate:  $t = -1.354$ ,  $df = 14$ ,  $p = 0.1$ ; slow rate:  $t = -0.479$ ,  $df = 14$ ,  $p = 0.32$ ).

### 3.3 Discussion

Our results for Scottish English did not replicate Van Lancker *et al.*'s foot-level polysyllabic shortening effect. No durational evidence for the word-internal foot hypothesis was obtained even at the slow speech rate, where constituent-related durational patterns would be expected to be amplified.

One interpretation of these results is that dialectal and lexical differences may play a role in the presence or absence of observable foot-related durational effects: Van Lancker *et al.*'s (1988) experiment tested American English speakers, while ours tested Scottish English speakers. It is possible that the different dialects of the participants yielded different results, or that foot-related effects only surface for particular types of materials. A pilot study of a word pair, *i*('*am*)(*bi*) vs. *i*('*ambus*), recorded by seven American English speakers at normal speech rate, did show results consistent with the foot hypothesis. This pair met our frequency criteria, and the durations of the second syllables of these words were roughly comparable. First syllable durations of these words were 3% longer in the two-foot condition, consistent with the foot hypothesis.<sup>18</sup> This result suggests the possibility that durational evidence for the foot might be found with different materials, different speakers or dialect groups.

Nevertheless, the comparison of effects in our foot experiment and those in our PWd experiments, both of which were conducted on data from the same Scottish English speakers, suggests that foot-related durational effects, if they exist at all, must be weaker than those of higher-level prosodic constituents such as PWd. In summary, not only phonological but also phonetic (durational) evidence for the foot is ambiguous at best.

<sup>18</sup> We also recorded the pair ('*rab*)(*bi*) vs. ('*rabbit*), and obtained the result that the rhyme duration of the former was 5% longer than the latter, which is also consistent with the foot hypothesis. However, this pair did not meet our frequency criteria, i.e. the former was less frequent than the latter.

## 4 General summary

We investigated whether differences in (a) word-internal morphological structure and (b) the lexical stress patterns of words are reflected in prosodic constituent structure. The outcome of our experiments on Scottish English was very different for these two types of structures, despite the fact that the same speakers participated in the two sets of experiments.

We found evidence for different prosodic constituent structure for Level II suffixed words as compared to monomorphemic words, and for Level II suffixed words as compared to Level I suffixed words. Rhyme durations of Level II stems were on average 4 to 6% longer than corresponding sequences in monomorphemic forms, and on average 7 to 8% longer than corresponding Level I suffixation cases in slow speech. However, no reliable durational differences were found for the comparison of double stressed words (i.e. monosyllabic foot cases) and single stressed words (i.e. disyllabic foot cases): rhyme durations of monosyllabic foot cases were not longer than those of disyllabic foot cases.

These results are consistent with the view that the morphological boundary between a stem and a Level II suffix is encoded as a prosodic word boundary. In contrast, none of our results from the same Scottish English speakers support the hypothesis that different lexical stress patterns are reflected in word-internal foot constituent-structure differences. As mentioned in §3.3 above, however, the null effect found in our experiments for the word-internal foot does not exclude the possibility that durational correlates of this constituent might be found for different dialects, lexical items and speakers. Further study of these effects is warranted.

## Appendix: Test materials

Target words are italicised here for clarity. They were not italicised in the printed scripts presented to speakers

### 1 Test materials used in Experiment 1

Set 1

- Pair 1 a. Two things he does with paper. Say 'tears paper'. Say '*tacks* paper'.
- b. Two kinds of paper. Say 'bank paper'. Say '*tax* paper'.
- Pair 2 a. Two things he does with corners. Say 'folds corners'. Say '*tucks* corners'.
- b. Two kinds of parties. Say 'cake party'. Say '*tux* party'.
- Pair 3 a. Two things he does with rain. Say 'drinks rain'. Say '*laps* rain'.
- b. Two kinds of rates. Say 'speech rate'. Say '*lapse* rate'.

Set 2

- Pair 1 a. Two things people felt about rain. Say 'loved rain'. Say 'missed rain'.  
b. Two kinds of rain. Say 'storm rain'. Say 'mist rain'.
- Pair 2 a. Two things he did with rolls. Say 'baked rolls'. Say 'passed rolls'.  
b. Two kinds of rolls. Say 'fresh rolls'. Say 'past rolls'.
- Pair 3 a. Two things people did with rails. Say 'nailed rails'. Say 'paced rails'.  
b. Two things people do with fence rails. Say 'cut rails'. Say 'paste rails'.

Set 3

- Pair 1 a. A pan used in the kitchen. Say 'baking pan' for me.  
b. A pan for cooking meat. Say 'bacon pan' for me.
- Pair 2 a. A special event at the cooking institute. Say 'canning event' for me.  
b. The 1 o'clock event at the Edinburgh Castle. Say 'cannon event' for me.
- Pair 3 a. A substitute for baking soda. Type 'raising substitute' for me.  
b. A substitute for dried fruit. Type 'raisin substitute' for me.
- Pair 4 a. An elementary mathematics exercise of height. Say 'summing height' for me.  
b. Height that matters for alpinists. Say 'summit height' for me.
- Pair 5 a. A naughty student's problem. Say 'socking problem' for me.  
b. A problem with an electrical hardware. Say 'socket problem' for me.
- Pair 6 a. To smoke a pipe. Say 'puffing' for me.  
b. A North Atlantic sea bird. Say 'puffin' for me.
- Pair 7 a. To move up and down. Say 'bobbing' for me.  
b. A cone holding thread. Say 'bobbin' for me.

Set 4

- Pair 1 a. Tools with a blade. Type 'axes' for me.  
b. A reference line. Type 'axis' for me.
- Pair 2 a. A clump of shrubs. Say 'bushes' for me.  
b. Equal to eight gallons. Say 'bushel' for me.

## **2 Test materials used in Experiment 2**

- Pair 1 a. A favourable attitude. Type 'acceptance' for me.  
b. He responds favourably. Type 'accepting' for me.
- Pair 2 a. Regular presence in class. Type 'attendance' for me.  
b. He goes to classes. Type 'attending' for me.
- Pair 3 a. The state of becoming gradually visible. Type 'emergence' for me.  
b. It becomes gradually visible. Type 'emerging' for me.
- Pair 4 a. Stubborn. Type 'insistent' for me.  
b. He states forcefully. Type 'insisting' for me.
- Pair 5 a. Relating to a group of words. Say 'phrasal' for me.  
b. He puts words together. Say 'phrasing' for me.

- Pair 6 a. A word for 'help'. Type 'assistance' for me.  
 b. He helps. Type 'assisting' for me.
- Pair 7 a. A word for 'keeping away'. Type 'avoidance' for me.  
 b. He keeps his distance. Type 'avoiding' for me.

### 3 Test materials used in Experiment 3

- Pair 1 a. A supersonic jet. Say 'Concorde' for me.  
 b. To overcome and take control. Say 'conquer' for me.
- Pair 2 a. A group of people living together. Say 'commune' for me.  
 b. Shared by two or more people. Say 'common' for me.
- Pair 3 a. A peak. Say 'climax' for me.  
 b. General weather conditions. Say 'climate' for me.
- Pair 4 a. An imaginary creature. Say 'centaur' for me.  
 b. The middle. Say 'centre' for me.
- Pair 5 a. Food for squirrels. Type 'acorn' for me.  
 b. A unit of land area. Type 'acre' for me.
- Pair 6 a. A surprise attack. Type 'ambush' for me.  
 b. A honey-yellow colour. Type 'amber' for me.
- Pair 7 a. A catalogue shopping store. Type 'Argos' for me.  
 b. A figure in Greek mythology. Type 'Argus' for me.

### REFERENCES

- Aronoff, Mark & S. N. Sridhar (1983). Morphological levels in English and Kannada or atarizing Reagan. In John F. Richardson, Mitchell Marks & Amy Chukerman (eds.) *Papers from the parasession on the interplay of phonology, morphology, and syntax*. Chicago: Chicago Linguistic Society. 3–16.
- Aylett, Matthew & Alice Turk (2004). The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47. 31–56.
- Baayen, R. Harald, Richard Piepenbrock & Hedderik van Rijn (1993). *The CELEX lexical data base*. [CD-ROM.] Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Beckman, Mary E. & Jan Edwards (1990). Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 152–178.
- Bell, Alan, Michelle L. Gregory, Jason M. Brenier, Daniel Jurafsky, Ayako Ikeno & Cynthia Girand (2002). Which predictability measures affect content word durations? In *Proceedings of Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology (PMLA)*. 1–5.
- Benua, Laura (1997). Affix classes are defined by faithfulness. *University of Maryland Working Papers in Linguistics* 5. 1–26.
- Bermúdez-Otero, Ricardo (1999). *Constraint interaction in language change: quantity in English and Germanic*. PhD dissertation, University of Manchester & University of Santiago de Compostela.

- Borowsky, Toni (1993). On the Word level. In Sharon Hargus & Ellen M. Kaisse (eds.) *Studies in Lexical Phonology*. San Diego: Academic Press. 199–234.
- Cambier-Langeveld, Tina (2000). *Temporal making of accents and boundaries*. PhD dissertation, University of Leiden.
- Campbell, N. W. & S. D. Isard (1991). Segment durations in a syllable frame. *JPh* **19**, 34–47.
- Chen, Matthew Y. (1990). What must phonology know about syntax? In Sharon Inkelas & Draga Zec (eds.) *The phonology–syntax connection*. Chicago: University of Chicago Press. 19–46.
- Cho, Taehong (2001). Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica* **58**, 129–162.
- Cho, Taehong & Patricia A. Keating (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *JPh* **29**, 155–190.
- Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.
- Christie, W. M. (1977). Some multiple cues for juncture in English. *General Linguistics* **17**, 212–222.
- Cohn, Abigail & John J. McCarthy (1994). Alignment and parallelism in Indonesian phonology. Ms, Cornell University & University of Massachusetts, Amherst.
- Cooper, William E. & Jeanne Paccia-Cooper (1980). *Syntax and speech*. Cambridge, Mass.: Harvard University Press.
- Cutler, Anne & Sally Butterfield (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory and Language* **31**, 218–236.
- Demuth, Katherine (1996). The prosodic structure of early words. In James L. Morgan & Katherine Demuth (eds.) *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Erlbaum. 171–184.
- Ferreira, Fernanda (1993). Creation of prosody during sentence production. *Psychological Review* **100**, 233–253.
- Fougeron, Cécile & Patricia A. Keating (1997). Articulatory strengthening at edges of prosodic domains. *JASA* **101**, 3728–3740.
- Fowler, Carol A. (1981). A relationship between coarticulation and compensatory shortening. *Phonetica* **38**, 35–50.
- Gerken, LouAnn (1994a). Young children's representation of prosodic structure: evidence from English-speakers' weak syllable productions. *Journal of Memory and Language* **33**, 19–38.
- Gerken, LouAnn (1994b). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language* **21**, 565–584.
- Hall, T. Alan (2001). The distribution of superheavy syllables in Modern English. *Folia Linguistica* **35**, 399–442.
- Halle, Morris & K. P. Mohanan (1985). Segmental phonology of Modern English. *LI* **16**, 57–116.
- Hammond, Michael (1999). *The phonology of English: a prosodic optimality-theoretic approach*. Oxford: Oxford University Press.
- Hayes, Bruce (1980). *A metrical theory of stress rules*. PhD dissertation, MIT. Published 1985, New York: Garland.
- Hayes, Bruce (1989). The prosodic hierarchy in meter. In Paul Kiparsky & Gilbert Youmans (eds.) *Rhythm and meter*. San Diego: Academic Press. 201–260.
- Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.
- Huggins, A. W. F. (1975). On isochrony and syntax. In G. Fant & M. A. A. Tatham (eds.) *Auditory analysis and perception of speech*. London: Academic Press. 455–464.
- Hulst, Harry van der & Norval S. H. Smith (eds.) (1982). *The structure of phonological representations*. 2 parts. Dordrecht: Foris.

- Inkelas, Sharon (1990). *Prosodic constituency in the lexicon*. New York: Garland.
- Itô, Junko & Armin Mester (1994). Reflections on CodaCond and Alignment. In Jason Merchant, Jaye Padgett & Rachel Walker (eds.) *Phonology at Santa Cruz 3*. Santa Cruz: Linguistics Research Center. 27–46.
- Kahn, Daniel (1976). *Syllable-based generalizations in English phonology*. PhD dissertation, MIT. Published 1980, New York: Garland.
- Kiparsky, Paul (1982). From cyclic to lexical phonology. In van der Hulst & Smith (1982: Part 1). 131–175.
- Kiparsky, Paul (2000). Opacity and cyclicity. *The Linguistic Review* 17. 351–365.
- Kiparsky, Paul (2008). Fenno-Swedish quantity: contrast in Stratal OT. In Bert Vaux & Andrew Nevins (eds.) *Rules, constraints, and phonological phenomena*. Oxford: Oxford University Press. 185–219.
- Ladd, D. Robert (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* 3. 311–340.
- Lehiste, Ilse (1960). An acoustic-phonetic study of internal open juncture. *Phonetica* 5. Suppl. 5–54.
- Lehiste, Ilse (1972). The timing of utterances and linguistic boundaries. *JASA* 51. 2018–2024.
- Liberman, Mark & Alan Prince (1977). On stress and linguistic rhythm. *LI* 8. 249–336.
- McCarthy, John J. (1982). Prosodic structure and expletive infixation. *Lg* 58. 574–590.
- McCarthy, John J. & Alan Prince (1993a). Generalized alignment. *Yearbook of Morphology 1993*. 79–153.
- McCarthy, John J. & Alan Prince (1993b). *Prosodic morphology I: constraint interaction and satisfaction*. Ms, University of Massachusetts & Rutgers University.
- Marshall, Jonathan (2004). *Language change and sociolinguistics: rethinking social networks*. Basingstoke: Palgrave Macmillan.
- Mester, Armin (1994). The quantitative trochee in Latin. *NLLT* 12. 1–61.
- Munson, Benjamin (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research* 44. 778–792.
- Nespor, Marina & Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
- Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen (2005). Lexical frequency and acoustic reduction in spoken Dutch. *JASA* 118. 2561–2569.
- Price, P. J., M. Ostendorf, S. Shattuck-Hufnagel & C. Fong (1991). The use of prosody in syntactic disambiguation. *JASA* 90. 2956–2970.
- Raffelsiefen, Renate (2005). Paradigm uniformity effects versus boundary effects. In Laura J. Downing, T. Alan Hall & Renate Raffelsiefen (eds.) *Paradigms in phonological theory*. Oxford: Oxford University Press. 211–262.
- Rakerd, Brad, William Sennett & Carol A. Fowler (1987). Domain-final lengthening and foot-level shortening in spoken English. *Phonetica* 44. 147–155.
- Redford, Melissa A. & Patrick Randall (2005). The role of juncture cues and phonological knowledge in English syllabification judgments. *JPh* 33. 27–46.
- Schwarzlose, Rebecca & Ann R. Bradlow (2001). What happens to segment durations at the end of a word? *JASA* 109. 2292.
- Scobbie, James M., Nigel Hewlett & Alice Turk (1999). Standard English in Edinburgh and Glasgow: the Scottish vowel length rule revealed. In Paul Foulkes & Gerard J. Docherty (eds.) *Urban voices: accent studies in the British Isles*. Oxford: Oxford University Press. 230–245.
- Scobbie, James M., Olga B. Gordeeva & Ben Matthews (2006). Acquisition of Scottish English phonology: an overview. *QMUC Speech Science Research Centre Working Paper WP-7*. Edinburgh: Queen Margaret University College.
- Selkirk, Elisabeth (1980a). The role of prosodic categories in English word stress. *LI* 11. 563–605.



- Selkirk, Elisabeth (1980b). Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff & Mary-Louise Kean (eds.) *Juncture*. Saratoga: Anna Libri. 107–129.
- Selkirk, Elisabeth (1981). On prosodic structure and its relation to syntactic structure. In Thorstein Fretheim (ed.) *Nordic Prosody II*. Trondheim: Tapir. 111–140.
- Selkirk, Elisabeth (1982a). *The syntax of words*. Cambridge, Mass.: MIT Press.
- Selkirk, Elisabeth (1982b). The syllable. In van der Hulst & Smith (1982: part 2). 337–383.
- Selkirk, Elisabeth (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass: MIT Press.
- Selkirk, Elisabeth (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3. 371–405.
- Selkirk, Elisabeth (1995). The prosodic structure of function words. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst: GLSA. 439–469.
- Siegel, Dorothy (1974). *Topics in English morphology*. PhD thesis, MIT. Published 1979, New York: Garland.
- Smith, Jennifer (2002). *Phonological augmentation in prominent positions*. PhD dissertation, University of Massachusetts, Amherst.
- Sproat, Richard (1993). Looking into words. In Sharon Hargus & Ellen M. Kaisse (eds.) *Studies in lexical phonology*. San Diego: Academic Press. 173–195.
- Sproat, Richard & Osamu Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *JPh* 21. 291–311.
- Steriade, Donca (2000). Paradigm uniformity and the phonetics–phonology boundary. In Michael B. Broe & Janet B. Pierrehumbert (eds.) *Papers in laboratory phonology V: acquisition and the lexicon*. Cambridge: Cambridge University Press. 313–334.
- Sugahara, Mariko & Alice Turk (2004). Phonetic reflexes of morphological boundaries at a normal speech rate. In Bernard Bel & Isabelle Marlien (eds.) *Proceedings of Speech Prosody 2004*. 353–356.
- Szpyra, Jolanta (1989). *The phonology–morphology interface: cycles, levels and words*. London & New York: Routledge.
- Treiman, Rebecca & Marie Cassar (1997). Can children and adults focus on sound as opposed spelling in a phoneme counting task? *Developmental Psychology* 33. 771–780.
- Turk, Alice (1993). *Effects of position-in-syllable and stress on consonant articulation*. PhD dissertation, Cornell University.
- Turk, Alice (1994). Articulatory phonetic clues to syllable affiliation: gestural characteristics of bilabial stops. In Patricia A. Keating (ed.) *Phonological structure and phonetic form: papers in laboratory phonology III*. Cambridge: Cambridge University Press. 107–135.
- Turk, Alice, Satsuko Nakai & Mariko Sugahara (2006). Acoustic segment durations in prosodic research: a practical guide. In Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.) *Methods in empirical prosody research*. Berlin & New York: de Gruyter. 1–28.
- Turk, Alice & James R. Sawusch (1997). The domain of accentual lengthening in American English. *JPh* 25. 25–41.
- Turk, Alice & Stefanie Shattuck-Hufnagel (2000). Word-boundary-related duration patterns in English. *JPh* 28. 397–440.
- Turk, Alice & Laurence White (1999). Structural influences on accentual lengthening in English. *JPh* 27. 171–206.
- Van Lancker, Diana, Jody Kreiman & Dwight Bolinger (1988). Anticipatory lengthening. *JPh* 16. 339–347.

- Waals, Juliette (1999). *An experimental view of the Dutch syllable*. PhD dissertation, Utrecht University.
- Walsh, Thomas & Frank Parker (1983). The duration of morphemic and non-morphemic /s/ in English. *JPh* **11**. 201–206.
- Warner, Natasha, Allard Jongman, Joan Sereno & Rachèl Kemps (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *JPh* **32**. 251–276.
- White, Laurence (2002). *English speech timing: a domain and locus approach*. PhD dissertation, University of Edinburgh.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA* **91**. 1707–1717.
- Wijnen, Frank, Evelien Krikhaar & Els den Os (1994). The (non)realization of unstressed elements in children's utterances: evidence for a rhythmic constraint. *Journal of Child Language* **21**. 59–83.